REVIEW ARTICLE

Accepted 27 July 2002
Revised 15 July 2002
Received 9 November 2001
Indian J. Pharm. Sci., 2002, 64(6): 509-514

Artificial Intelligence in Pharmaceutical Research: Theory and Applications

P. MOUDGIL, A. K. TIWARY AND S. GUPTA*
Department of Pharmaceutical Sciences and Drug Research,
Punjabi University, Patiala-147 002.

Artificial intelligence includes methods, tools and systems devoted to simulate human brain. The brain activity comprises of logical-cum-inductive knowledge acquisition and reasoning for solving problems. The future seems to be closely tied to research programs in artificial intelligence. The theory of neural network operation and potential applications of neural network based artificial intelligence in pharmaceutical sciences are reviewed in the present article.

Artificial intelligence is the study of behavior and implementation of a program that exhibits intelligent behavior. There are two main categories of artificial intelligence developments. The first includes methods and systems that simulate human experience and draw conclusions from a set of rules, like expert systems. The second includes systems that model the way the brain works, like artificial neural networks[1]. Expert systems consist of rules constructed by human experts that have been translated into a computer program. The advantage of an expert system is that reasons for a given interpretation are known. The expert systems are, however, not good at pattern recognition[2]. Neural networks are learning machines that analyze systems and seek to establish mathematical relationships between seemingly unrelated variables. The diverse applications of neural networks can be summarized into pattern recognition, predictive purposes and modeling in a variety of disciplines like biological science, engineering and business[3-4].

ARTIFICIAL INTELLIGENCE

In 1970's, it began to be realized that a computer program could display intelligent behavior if the domain it deals with is sufficiently narrowed. The transformation from logic-oriented to knowledge-based approach is a new transformation in the field of artificial intelligence. The emergence of intelligence is based on two computational points of views, one based on symbolism and the other based on connectionism. The former approach models intelligence using symbols while the latter approach model intelligence using connections and associated weights. However, most of the knowledge can only be represented symbolically[5]. In contrast to symbolic approach, the neural network approach adopts brain metaphor suggesting that intelligence emerges through a large number of neurons connected together, each performing simple computation. The long-term knowledge of a neural network is encoded as a set of weights on connections between units. The neural network architecture has therefore been dubbed as 'the connectionist'[6].

The Japanese, in 1981, were first to demonstrate their commitment to artificial intelligence. They launched a research and development program in artificial intelligence called Fifth Generation. In 1983, to develop advanced technologies that apply artificial intelligence technique, a consortium of private companies was formed in U.S.A. Subsequently other countries slowly and gradually initiated plans in the field of artificial intelligence.

Artificial Neural Networks (ANNs) versus human brain:

The ANN is a biologically inspired system designed to simulate neurological processing ability of human brain. The human brain is most highly organized matter in universe. Human brain consists of about 100 billion neurons and about 100,000 times that many neuron connections. Neuron consists of a cell body, hub of which is the soma, inputs are the

*For correspondence
E-mail: gsunil70@rediffmail.com

dendrites and output is the axon. When an axon connects to a dendrite of another neuron, the connection is called a synapse. Neurons are organized in a fully connected network and act like a messenger in receiving and sending impulses. The result is an intelligent brain capable of learning, prediction and recognition. It can recognize a face in few milliseconds and adjust eyesight several orders of magnitude. It can take note of many things happening at the same time (parallel processing). The brain is an excellent pattern recognition tool. The ANNs mimic working of human brain. The ANN system consists of simple neurons that operate in parallel. The function of neurons is determined by network structure, connection strength and processing carried out by neurons. The ANNs learn (or are trained) through experience and not by programming. They gather know-how by detecting patterns and relationships in the data. The ANNs seem to fulfill cherished dream of scientists to develop machines that can think like human beings.

Out of numerous neural network architectures[7] developed so far, error back-propagation network developed by Rumelhart *et al.*[8] is most widely and successfully applied architecture. The error back-propagation network consists of input layer, output layer and one or more hidden layers. The input layer provides data from the external world. The mapping of input data by neural network into interpretable results is done by a representative signal generated by output layer. The activation of hidden layers represents an internal encoding by neural network for solution of the given problem using an intermediate level of abstraction. The hidden layers are called so because these layers are not directly accessible; they are concealed behind another layer in each direction. The ability of neural networks to classify information separated by non-linear boundaries depends on hidden layers.

The building component of ANN, called neuron, receives many signals as weighted process variables from the response of other units. The signals are multiplied by connection weights and then combined to produce a composite signal. The mapping of composite input via some activation function then determines activation-state of the receiving neuron. The activation function is weighted sum of neuron inputs. The signal is then passed through a transfer function to produce response for that neuron. The most commonly used transfer function is sigmoid function shown in fig. 1. The advantage of sigmoid function is that it can accommodate large signals without saturation while allowing small signals to pass without excessive attenuation.

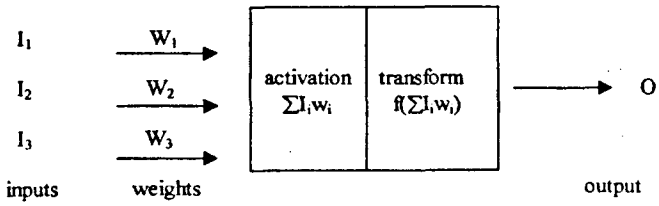

Fig. 1: Operation of a transfer function.

I_1, I_2, I_3 = process variables, W_1, W_2, W_3 = associated weights and O = response.

The application of back-propagation technique to a problem requires three simple steps-network design, learning or training and usage[9]. In the first step, network design consists of finalizing the number of layers and neurons in each layer. The number of process variables and response units are determined by the data of problem. The network is made to learn or is trained in second step. The knowledge possessed by network is represented through number of interconnections and their weights that neural network uses for learning or training. The learning through weight adjustment can be supervised or unsupervised. On the other hand training consists of feed forward step and feed back step. Supervised learning results when a network is repeatedly presented with an input pattern and a desired output response. If 'I' is input pattern and 'R_D' is desired response pattern, weights are adjusted to achieve output response 'R_O' within the user-defined limit of desired response 'R_D', that is until 'R_O-R_D' achieves error goal. When error goal is near zero, training process terminates at a point where neural network produces correct response for given input patterns. In unsupervised learning, no desired response 'R_D' is available to guide the system. The learning is through input pattern 'I' alone. The neural network system itself then decides features to be used for grouping input data. This process is also referred to as self-organization or adaptation[10]. The feed forward step begins with presentation of process variable pattern and continues through activation levels to propagate through hidden layers. In each hidden layer, processing unit sums the input and applies sigmoid function to compute its response as shown in fig. 2. In feed back step, error values are calculated for all processing units and weight changes are calculated for all interconnections. The calculations start at output layer and progress backward through the network to input layer. Thus, each neuron has one additional weight as an input that allows an additional degree of freedom when trying to minimize training error. After network design is completed and system is trained, the network is at third step-ready for use on new patterns.

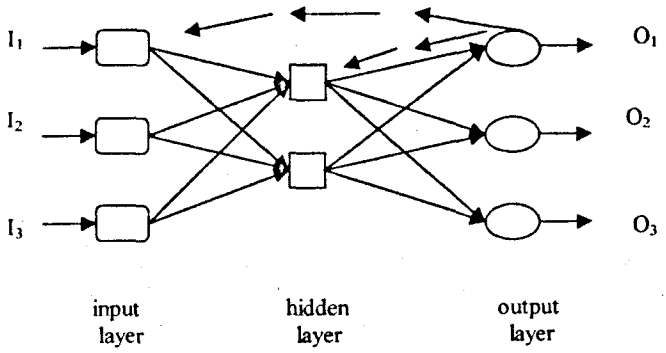

Fig. 2: A three layered artificial neural network.

I_1, I_2, I_3 = process variables and O = response.

APPLICATIONS IN PHARMACEUTICAL SCIENCES

ANNs are machine based computational techniques that attempt to model neurological processing ability of the human brain. The computations based on neural network are different from standard statistical method of analysis. Any problem that requires control and sensor processing is suited for application of ANN technology. The ANNs require less formal statistical training. The neural networks can detect complex non-linear relationships by using multiple training algorithms. To solve problems the ANNs can combine literature based and experimentally generated data. The neural networks can function as autoassociation model, heteroassociation model, classification model, optimization model and self-organization model depending on the nature of application[5]. The Hopfield type neural networks are autoassociation models that retrieve memory based on part of memory itself. The heteroassociation models retrieve memory in one set using another memory in a different set. Classification models, according to the response activation, group a given object presented to neural networks. For binary outputs 0 corresponds to one class and 1 corresponds to another class while for continuous output between 0 and 1 the decision is made using a threshold value of 0.5. The Boltzman machines and Hopfield type nets are optimization models that solve combinatorial problems. The combinatorial optimization problems often lack efficient solutions on a digital computer. The Kohonen type networks are self-organization models that learn and organize information without being given correct answers for process variable patterns.

Pattern recognition and Interpretation of analytical data:

The neural networks can recognize patterns from a complex analytical data. When spectrum of unknown is a superposition of known spectra the ANNs have been employed to determine the composition of unknown sample. Whole of the spectrum is considered in identification process instead of individual peaks. The only concern is involvement of time span for training neural networks. The ranitidine hydrochloride tablet is a multi-component tablet formulation in which there is significant overlap of spectral pattern of ingredients. Agatonovic-Kustrin *et al.*[11] have used solid state techniques, diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) and X-ray powder diffractometry (XRPD) to analyze polymorphic purity of ranitidine hydrochloride. The DRIFT spectra and XRD powder patterns of different mixtures of Form 1 and Form 2 of ranitidine hydrochloride have been analyzed by using ANNs. A neural network consisting of three layers of neurons was trained using a back-propagation learning rule to recognize specific patterns of individual constituents of formulations from the overall spectral pattern. The classification network identified and quantified all components in tablet when exposed to complex formulation containing only form F1 crystals. A method has been developed by Sondergaard *et al.*[12] to classify crossed immunoelectrophoretic patterns using ANNs. It has been found that ability to generalize was increased by the addition of noise to the input patterns during training. The trained neural network classified distorted patterns correctly within an error range of 1%. Madden *et al.* have used ANN as the basis of computer-assisted optimization method for selection of optimal gradient conditions for anion separations. The ANNs with 1-10-9 architecture has been found to be rapid and accurate in predicting retention times for anions in linear gradient elution ion chromatography with hydroxide eluents[13]. Lee and co workers have used ANNs to analyze *in vivo* magnetic resonance spectra of breast cancer patients[14]. RESCUE, an ANN tool has been created for NMR spectral assignment of proteins. RESCUE determines the type of amino acid from chemical shift values observed in ^1H spectrum[15].

Formulation and optimization:

The initial step in development of drug delivery system is design of pilot plant experiments that involve extensive experimentation. The simulation by neural networks, using previously collected data, can be utilized to identify right formulation and processing parameters. The network is updated as new results keep pouring in and hence its accuracy increases. This assures that formulator is always on the right track. The optimum *in vitro* drug release profile from controlled release hydrophilic matrix capsules predicted by using delta back-propagation neural network and response surface methodology has been compared by Hussain *et al.*[16] The predictions using ANNs were more accurate than those

using response surface methodology. To design pharmaceutical formulations a formulator often comes across situation when any one property of formulation can be maximized and remaining desired characteristics has to be compromised. The ANNs are useful in such multi-objective simultaneous optimization problems. The reliability of ANNs in optimizing ketoprofen hydrogel ointment has been demonstrated by Takayama *et al.* against classical response surface methodology approach[17]. Wu *et al.* prepared different osmotic pump tablets of salbutamol sulfate[18]. The average drug release rate for first 8 h and correlation of cumulative amount of drug released with time was used to characterize the release of salbutamol. The release parameters and causal factors such as amount of hydroxypropyl methylcellulose, polyethylene glycol 1500 in coating solution and coat weight were used to train the ANN. Both training and test sets of data were exposed to ANN. The training process of ANN was terminated when a satisfactory value of error function for test data was obtained. The release profile of the optimum formulation was in agreement with results predicted by ANN.

Structure prediction and classification of proteins:

Being a computational analysis technique, ANNs are suitable for recognition of domains, classification of proteins, prediction of enzyme class, sequence analysis of DNA/RNA and proteins. A feed forward ANN consisting of six input and six hidden units with sigmoid transfer function has been used for recognition of domains in protein sequence by Murval *et al.*[19] A preliminary step in protein fold recognition and ab *initio* structure prediction is the classification of proteins. Pasquier *et al.* used a hierarchical network named PRED-CLASS to classify proteins into four classes such as transmembrane, fibrous, globular and mixed proteins[20]. The PRED-CLASS trained using 50 protein sequences, correctly predicted 371 out of a set of 387 proteins with an accuracy of 96 percent. The ANN based algorithms has been used to identify, characterize and predict stabilization center elements from primary structure of single proteins and amino acid sequences of homologous proteins[21]. The stabilization center elements present in proteins stabilize protein structures by preventing their decay. The prediction of long chain fatty acid transport protein FadL topology[22], prediction of secondary structures of clostridial neuroprotein-C fragment[23], DNA/RNA and protein sequence analysis[24] are other applications where ANN technology has been exploited. The development of efficient delivery systems for protein drugs that are structurally complex and fragile is a tedious and challenging task because proteins have specific sites responsible for different processes. Variations at these specific sites lead to loss of therapeutic action. The neural networks have been employed to predict eukaryotic protein phosphorylation sites[25], to recognize active sites and to predict enzyme class with high accuracy for novel protein structures[26].

QSAR, QSPR and molecular modeling:

Quantitative structure-activity relationships (QSAR) correlate structural or property descriptors of compounds with chemical or biological activities. All QSAR studies are based on the fundamental concept of interdependence of biological activities on physicochemical parameters. The physicochemical descriptors and topological parameters can be determined by computational methods. To predict anti-HIV activity of 1-[2-(hydroxyethoxy)methyl]-6-(phenylthio)thymme (HEPT) derivatives QSAR has been established using multiple linear regression (MLR) and a fully connected feed forward model with 6-6-1 ANN architecture. The mean square error for prediction set using MLR was twice as compared to error obtained using ANN[27]. The superiority of ANNs over MLR for establishing QSAR has been claimed by Wiese and Schaper[28], Tetko *et al.*[29] and Yamamura *et al.*[30] To predict antimicrobial property of quinolone derivatives on the basis of their chemical structures a new topological method has been developed by Jaen-Oltra *et al.*[31] A neural network, with suitable set of topological descriptors and training algorithms, was used to determine the minimum inhibitory concentration of quinolones. The predicted values were in agreement with experimentally observed values. Huuskonen *et al.* predicted aqueous solubility within a series of structurally related drugs using neural network models[32]. Topological descriptors were used to establish relationship between structures of compounds with their aqueous solubility. To predict aqueous solubility of a congeneric set of compounds useful models containing simple structural parameters were obtained using neural networks. Dinner *et al.*[33] have developed a quantitative structure-property relationship (QSPR) model that employs a genetic algorithm to pick attributes for folding and an ANN to derive functional dependence of folding on the sequence attributes. Despite problems of large training time neural nets have outperformed conventional techniques in QSAR studies.

Epidemiology:

The community pharmacists have applied basic principles of epidemiology or medical ecology to determine occurrence of diseases in defined populations based on interaction of the host and environment. To predict occurrence of diseases modeling approaches have been used. The ANN

has been employed to group 39 European countries according to 'Health for All' database, maintained by WHO, with respect to life expectancy, infant mortality rate, standardized death rates for diseases of circulatory system, external causes of injury and poisonings[34]. A network for analysis of alcoholism data[35] has been designed but such approaches have their limitations as large populations are under study, which pose difficulties in initial identification and finally quantification of causative variables.

Biopharmaceutics:

Bioavailability is a measure of fraction of drug substance reaching the systemic circulation to elicit therapeutic effect. All drug regulatory authorities around the world require bioavailability or bioequivalence data from the manufacturers to justify their claims. The *in vitro-in vivo* correlations (IVIVC) are of great interest for pharmaceutical industry to avoid bioequivalence studies that are predicted to produce negative results. The neural nets being good at pattern recognition task have been used to establish correlation between *in vitro* and *in vivo* performance of formulations. Dowell *et al.* have developed a number of unique ANN configurations to predict IVIVC from different formulations of same product[36]. The configuration variables included architectural structure, learning algorithm and input-output association structures.

Pharmacokinetics and pharmacodynamics:

The ANN technology offers an exciting alternative to monitor complex kinetic interactions between drug substances and physiological system that are usually monitored by pharmacodynamics. Veng Pederson and Modi added a new perspective to pharmacodynamic modeling by introducing the concept of neural network modeling[37]. Gobburu and Shelver used neural networks to predict the pharmacokinetic properties of β-adrenoreceptor antagonists in humans and concluded that neural networks can be a powerful tool in exploration of quantitative structure-pharmacokinetic relationships[38]. Gobburu and Chen have used neural networks for integrated pharmacokinetic and pharmacodynamic analysis[39]. To predict pharmacokinetic parameters in human beings Hussain *et al.* have developed a neural network using laboratory animal data[40]. The feedforward back-propagation network with a combination of physicochemical properties and animal pharmacokinetic parameters has been proposed to predict human pharmacokinetic parameters[41]. The use of ANN approach is growing rapidly for the analysis of population pharmacokinetic data monitoring of complex kinetic interactions in large populations. Chow and co-workers have applied neural network approach on the data collected from pediatric patients undergoing treatment with tobramycin for bacterial infection[42]. Neural nets were trained to find relationship of plasma tobramycin concentration with patient-related demographic factors and relationship of dosing regimens with time of sample collection. A standard population pharmacokinetic modeling program, NONMEM was compared with ANNs. The relationships predicted by neural network were similar to those predicted by NONMEM. The average errors between absolute and predicted values were found to be very low for neural networks. The strength of the neural networks is that they do not assume a specific model The learning ability of ANNs to establish the input and output relationships from the data provided to them simplifies the modeling work involved in traditional population pharmacokinetic data analysis.

Simulation of rat skin permeability:

To simulate rat skin permeability authors have formulated chitosan membranes. The membranes of varied composition were prepared to evaluate effect of critical process variables on permeation parameters such as flux and lag time. The permeability in chitosan membranes was predicted using model dependent and model independent techniques. The sum of square of residuals was less as compared to prediction made by multiple linear regression method when permeation parameters were predicted by 3-5-2 ANN architecture.

CONCLUSION

The ANNs are newly developed modeling strategies and an alternative to conventional modeling techniques. The ANNs need no special computer as neural nets are described using mathematical models and implemented using ordinary computer software. At present, training time for networks is long but the advantages offered by neural nets are overwhelming. Neural computing technology is capable of solving problems involving complex pattern recognition. This advantage has been utilized in pharmaceutical product development. The ANNs have an edge over response surface methodology because they allow incorporation of literature and experimental data to solve common problems in pharmaceutical industry. The use of artificial intelligence in pharmaceutical research is new but growing at a fast pace.

REFERENCES

1. Agatonovic-Kustrin, S. and Beresford, R., **J. Pharm. Biomed. Anal.**, 2000, 22, 717.
2. Holst, H., Astrom, K., Jarund, A., Palmer, J., Heyden, A., Kahl,

F., Tagil, K., Evander, E., Gunnar, S. and Edenbrandt, L., **Eur. J. Nucl. Med.,** 2000, 27, 400.
3. Furlong, J.W., Dupuy, M.E. and Heinsimer, J.A., **Amer. J. Clin. Pathol.,** 1991, 96, 134.
4. Baxt, W.G., **Lancet,** 1995, 346, 1135.
5. Fu, L.M., In; Neural Networks in Computer Intelligence, McGraw-Hill, Inc., Singapore, 1994, 3.
6. Feldman, J.A. and Ballard, D.H., **Cognitive Science,** 1982, 6, 205.
7. Simpson, P.K., In; Artificial Neural Systems, Pergamon Press, New York, 1990, 79.
8. Rumelhart, D.E., Hinton, G.E. and Williams, R.J., In; Parallel Distributed Processing: Explorations in the Microstructures of Cognition, Vol. 1, MIT Press, Cambridge, 1986, 318.
9. Achanta, A.S., Kowalski, J.G. and Rhodes, C.T., **Drug Develop. Ind. Pharm.,** 1995, 21, 119.
10. Patterson, D.W., In; Introduction to Artificial Intelligence and Expert Systems, Prentice Hall of India Pvt. Ltd., New Delhi, 2001, 345.
11. Agatonovic-Kustrin, S., Rades, T., Wu, W., Saville, D. and Tucker, I.G., **J. Pharm. Biomed. Anal.,** 2001, 25, 741.
12. Sondergaard, I., Krath, B.N. and Hagerup, M., **Electrophoresis,** 1992, 13, 411.
13. Madden, J.E., Avdalovic, N., Haddad, P.R. and Havel, J., **J. Chromatogr. A.,** 2001, 910, 173.
14. Lee, J.K., Tsai, S.C., Ho, Y.J., Chanclai, S.P. and Kao, C.H., **Anticancer Res.,** 2001, 21, 1481.
15. Pons, J.L. and Delsuc, M.A., **J. Biomol. NMR,** 1999, 15, 15.
16. Hussain, A.S., Yu, X. and Johnson, R.D., **Pharm. Res.,** 1991, 8, 1248.
17. Takayama, K., Fujikawa, M. and Nagai, T., **Pharm. Res.,** 1999, 16, 1.
18. Wu, T., Pan, W., Chen, J. and Zhang, R., **Drug. Develop. Ind. Pharm.,** 2000, 26, 211.
19. Murval, J., Vlahovicek, K., Szepesvari, C. and Pongor, S., **Genome Res.,** 2001, 11, 1410.
20. Pasquier, C., Promponas, V.J. and Hamordrakas, S.J., **Proteins,** 2001, 44, 361.
21. Dosztanyl, Z., Fiser, A. and Simon, I., **J. Mol. Biol.,** 1997, 272, 597.
22. Cristalli, G., DiRusso, C.C. and Black, P.N., **Arch. Biochem. Biophys.,** 2000, 377, 324.
23. Lebeda, F.J., Umland, T.C., Sax, M. and Olson, M.A., **Protein Chem.,** 1998, 17, 311.
24. Wu, C.H., **Comput. Chem.,** 1997, 21, 237.
25. Blom, N., Gammeltoft, S. and Burnak, S., **J. Mol. Biol.,** 1999, 294, 1351.
26. Stahl, M., Taroni, C. and Schneider, G., **Protein Engg.,** 2000, 13, 83.
27. Jalali-Heravi, M. and Parastar, F., **J. Chem. Inf. Comput. Sci.,** 2000, 40, 147.
28. Wiese, M. and Schaper, K.J., **SAR QSAR Environ. Res.,** 1993, 1, 137.
29. Tetko, I.V., Villa, A.E. and Livingstone, D.J., **J. Chem. Inf. Comput. Sci.,** 1996, 36, 794.
30. Yamamura, S., Nishizawa, K., Hirano, M., Momose, Y. and Kimura, A., **J. Pharm. Sci.,** 1998, 3, 95.
31. Jaen-Oltra, J., Salabert-Salvador, M.T., Garcia-March, F.J., Perez-Gimenez, F. and Tomas-Vert, F., **J. Med. Chem.,** 2000, 43, 1143.
32. Huuskonen, J., Sato, M. and Taskinen, J., **J. Pharm. Sci.,** 1997, 86, 450.
33. Dinner, A.R., So, S.S. and Karplus, M., **Proteins,** 1998, 33, 177.
34. Lloyd-Williams, M. and Williams, T.S., **Top Health Inf. Manage.,** 1996, 17, 26.
35. Li, W., Haghighi, F. and Falk, C.T., **Genet. Epidemiol.,** 1999, 17 Suppl 1, S223.
36. Dowell, J.A., Hussain, A., Devane, J. and Young, D., **J. Pharm. Sci.,** 1999, 88, 154.
37. Veng-Pederson, P. and Modi, N.B., **J. Pharmacokinet. Biopharm.,** 1992, 20, 397.
38. Gobburu, J.V. and Shelver, W.H., **J. Pharm. Sci.,** 1995, 84, 862.
39. Gobburu, J.V. and Chen, E.P., **J. Pharm. Sci.,** 1996, 85, 505.
40. Hussain, A.S., Johnson, R.D., Vachharajani, N.N. and Ritschel, W.A., **Pharm. Res.,** 1993, 10, 466.
41. Ritschel, W.A., Akileswaran, R. and Hussain, A.S., **Methods Find. Exp. Clin. Pharmacol.,** 1995, 17, 629.
42. Chow, H.H., Tolle, K.M., Roe, D.J., Elsberry, V. and Chen, H., **J. Pharm. Sci.,** 1997, 86, 840.