An overview of neural network applications

H.J. Kappen

Foundation for Neural Networks SNN,
Department of Biophysics, University of Nijmegen,
Geert Grooteplein 21, NL 6525 EZ Nijmegen, The Netherlands

Neural networks attempt to use human-like computing strategies to improve the performance of computers for artificial intelligence applications. A short introduction about neural networks and an overview of application areas is given.

1 Introduction

The development of computers has been very fast and computers have come to provide a huge amount of computing power. Nevertheless, in many respects computers still lack the flexibility of processing and are far behind humans. Examples are pattern recognition, scene interpretation, problem solving based on incomplete and contradictory information, and learning capability.

Neural computing [1], [2], [3] [4] is a relatively recent development in the information sciences, an outgrowth of artificial intelligence research in the 1950s and 1960s. Neural networks attempt to use human-like computing strategies to improve the performance of computers in these areas. Neural networks are so named because they exhibit certain analogies, at least superficially, to the way in which arrays of neurons probably function in biological learning and memory. They differ from the usual computer programs in that they "learn" from a set of examples rather than being programmed to get the right answer. The information is encoded in the strength of the network's "synaptic" connections.

Essential aspects of neural information processing are highly parallel execution of computation, integration of memory and process, as well as a performance which is robust against noise. Research in specific fields of application must combine knowledge from different research disciplines, such as computer science, physics, neuro-biology, psychology and engineering.

In the last few years the algorithms have advanced to the point where they can be succesfully applied to a large number of industrial and other practical applications. Fig. 1 shows the industrial interest in

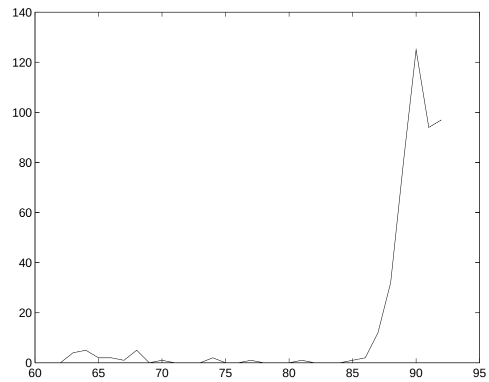

Fig. 1. European patents using neural networks from 1960 to 1992.

neural networks, expressed in number of patents. Note the first wave of interest in the 60s, as a result of the popular field of 'cybernetics' in the 50s. During this period, the first learning networks were designed. The idea that computers could learn, led to a large industrial interest. However, during the mid 60s it became apparent that these networks were not powerful enough for many practical problems. In addition, symbolic and rule based methods gained much popularity during this period.

During the 60s and 70s several more powerful neural network architectures and learning rules were proposed [5, 6]. These algorithms are in essence the now popular multi-layered perceptrons neural network with the back-propagation learning algorithm. However, the limited computing resources of the time did not allow to use these methods to solve practical problems. Since 1985, there is a marked increase in industrial activity in neural networks. This is largely due to the spectacular increase in computing power during the last decades and the gained insight that symbolic methods are not as powerful as expected for noisy real-world knowledge domains.

Some important application areas are time-series forecasting with financial application [7] [8], process control [9], robotics, pattern recognition [10], chemical analysis and database analysis. For a good overview of succesful applications of neural networks see [11].

2 Neural Networks

Most neural networks are used for classification problems or for function estimation. Such problems consist of a number of feature vectors, together with their desired class or function value. A typical neural network architecture is given in fig. 2. The fundamental building blocks are units which can be

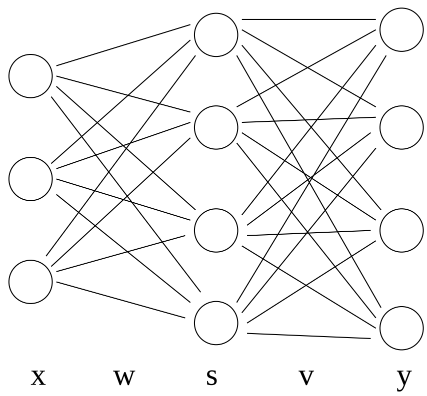

Fig. 2. A multi-layered perceptron with input layer x, hidden layer s and output layer y. Connections between layers are adapted during the training process.

likened to neurons, and weighted connections which can be likened to synapses. The networks have a number of input units in which the feature vectors are encoded. The output units encode the different classes or the function values. Between the inputs and outputs of the network, a hidden layer of neurons is present whose number of units can be varied. As shown in Fig. 2 each hidden unit is connected to all inputs and to all outputs. The more hidden units, the more complex the patterns that can be learned.

Using a set of feature vectors with their known output values, we can train the network. This training procedure is done using the "backpropagation algorithm" [1] for function estimation, or Boltzmann Machine learning classification problems [12]. Neural networks used in this way can be shown to be able to approximate any problem arbitrarily well, given enough hidden units [13]. Therefore, by increasing the number of hidden units, the network performance on the training data can be improved to arbitrary precision. However, the prediction of the network on independent test data, which is called the generalization performance, does not necessarily improve for larger networks. Therefore, the performance of the network is verified on independent test data. In this way the optimal number of hidden units is chosen.

The generalization capabilities of a neural network is generally good compared with other methods. However, generalization can be poor when the network is trained on insufficient data. Generalization can be significantly improved by introducing pruning or weight decay mechanisms or complexity reducing constraints [14].

The learning process itself consists of an optimization procedure in a complex high dimensional space. The optimal values of the network's synaptic connections, also called the weights, are obtained after learning and represent the networks knowledge of the task that it has learned. This optimization process is plagued by so-called local minima, which are sub-optimal solutions at which the learning process may terminate. In [15], a theoretical study led to learning procedures which facilitate the escape from local minima.

3 Some examples of applications

In this section some examples of succesful applications of neural networks are presented. These applications are taken from a list of approximately 100 casestudies that is being compiled during the Esprit project SIENA. The aims of SIENA are to assess the impact of neural networks on European industry. Some of these case studies were presented during the SNN conference on neural network applications in 1995 [11].

3.1 Prediction of newspaper sales

For "De Telegraaf", one of the major Dutch publishing companies, SNN developed a method to estimate the number of sales of newspapers and magazines that will be sold in supermarkets, bookshops, kiosks etc. The ideal is to deliver the number of copies that will be sold plus an additional one to verify that each customer has been able to buy it. This is, however, not realistic, as the sale of newspaper is a highly stochastic process. Thus, for each sales point a margin above the estimated number of sales must also be calculated.

The major difficulty was the required robustness of the method. It had to be useful for thousands of different sales points with largely different characteristics. The number of sales could range from a handful to several hundreds for the largest shops. The variability of sales could be small or large and might depend on season as well.

Results on historical data reveal that up till 40% of the sell outs can be avoided without increase of the number of unsold copies [16]. This method is currently tested by "de Telegraaf".

3.2 Current Prediction for Shipping Guidance in IJmuiden

The North Sea Directorate of Rijkswaterstaat provides, among other things, forecasts of hydrological parameters to Dutch port authorities.

The IJ-channel, that leads to the port of IJmuiden, 25 km west of Amsterdam, is accessible to ships with drafts up to 16.46 m. The relatively strong cross-channel current in front of the harbour moles has an important bearing on navigational safety of deep draft ships. To ensure a safe passage of the harbour moles a current criterion has been established. This criterion is exceeded in almost every tidal cycle.

A reliable current forecast is important to reduce waiting time for deep draft ships. Rijkswaterstaat, in collaboration with All Fours Neural Network Applications, trained a neural network for current prediction. The neural network has been trained with nine months of current measurements and simultaneous wind and water level data from several locations in the North Sea. The neural network is implemented on a PC in the Hydro Meteo Centre Rijnmond (HMR) at Hook of Holland. A 24 hour current forecast on the basis of on-line measurements and forecasts of wind and water level is provided four times a day. The current forecast is relayed to the port authority of IJmuiden. Since April 1994, the current prediction system is in use helping IJmuiden pilots to make their sailing plans [17].

3.3 Prediction of Yarn Properties in Chemical Process Technology

To produce yarns with the desired properties in an economic way, one need to know the relation between the available production technologies, molecular structures and the final yarn properties. These relations can be obtained by chemical-technological experiments. However these experiments are costly and time consuming. Insights in these relations could considerably save on the technological experiments. However, these relations are too complex for a quantitative judgement by human experts.

Akzo, in collaboration with the department of Analytical Chemistry of the University of Nijmegen, developed a neural networks application for this problem. To provide examples to train the neural network 295 yarns were produced with different structures and properties. Of each of the yarns, 5 structure parameters and 15 properties were determined. With these data, the network has been trained. The trained network is able to successfully predict the properties from the structures of new yarns. The neural network is now used within Akzo as a tool for the researchers to find out how the best yarns are synthesized [18].

3.4 Recognition of Exploitable Oil and Gas Wells

Whether economically recoverable oil or gas is present in subsurface reservoirs is highly dependent on physical properties of the reservoir rock such as porosity and permeability. These properties can be determined directly from actual rock samples (cores) taken from wells. For economical reasons, however, cores are only available from a limited number of wells. In most instances, formation analysts have to rely on measurements from wireline tools. The problem is the inherent variability of the wireline data, due to gradations in rock characteristics, effects of data acquisition and statistical fluctuations in radiation measurements. This makes the identification of rock types very difficult, in particular with complex and heterogeneous reservoirs.

Shell Research trained neural networks to classify rock types as function of the wireline data. For different geological environments different neural networks were trained. The desired outputs were provided by expert geologists. The degree of consensus between the neural network and its trainer were roughly equivalent to the degree of consensus among different geologists.

The neural network has been incorporated in Shell's geological computing environment, in which the formation analyst can train and apply the neural network via a user-friendly graphical interface. The advantage of this approach is that it enables the formation analysts to classify rock types from wireline data at reduced effort. This neural network performs significantly better than standard statistical techniques [19]

References

1. D. Rumelhart, G. Hinton, and R. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.
2. D. Ackley, G. Hinton, and T. Sejnowski. A learning algorithm for Boltzmann machines. *Cognitive Science*, 9:147–169, 1985.
3. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci. USA*, 79:2554–2558, 1982.
4. T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.
5. S. Amari. A theory of adaptive pattern classifiers. *IEEE Transactions on Electronic Computers*, 16:299–307, 1967.
6. A. Bryson, W. Denham, and S. Dreyfuss. Optimal programming problem with inequality contraints. i: Necessary conditions for extremal solutions. *AIAA Journal*, 1:25–44, 1963.
7. H. White. Economic proediction using neural networks: the case of ibm daily stock returns. In *Advances in Neural Processing Systems III*, pages 451–459. Morgan Kaufmann, 1988.
8. A.S. Weigend, B.A. Huberman, and D.E. Rumelhart. Predicting the future: A connectionist approach. *International Journal of Neural Systems*, 1:193–209, 1990.
9. P. Werbos. *Neural Networks for Control.* MIT Press, Cambridge, USA, 1990.
10. I.K. Sethi and A.K. Jain, editors. *Artificial Neural Networks and Statistical Pattern Recognition.* North-Holland, 1991.
11. B. Kappen and C. Gielen, editors. *Neural Networks: Artificial Intelligence and Industrial Applications. Proceedings of the Third Annual SNN Symposium on Neural Networks, Nijmegen, The Netherlands, 14-15 September 1995*, London, 1995. Springer.
12. H.J. Kappen. Deterministic learning rules for Boltzmann machines. *Neural Networks*, 8:537–548, 1995.
13. G. Cybenko. Approximation by superposition of a sigmoidal function. *Math. Control Signals Systems, Vol. 2*, pages 303–314, 1989.
14. S. Haykin. *Neural Networks. A Comprehensive Foundation.* Macmillan College Publishing Company, New York, 1994.
15. T. Heskes, E. Slijpen, and B. Kappen. Learning in neural networks with local minima. *Physical Review A*, 46:5221–5231, 1992.
16. M. Theeuwen and H.J. Kappen. Neurale netwerken en voorspelling losse verkoop. confidential, 1995.
17. J.C. Wüst. Current prediction for shipping guidance. In Bert Kappen and Stan Gielen, editors, *Neural networks: artificial intelligence and industrial applications Proceedings of the 3rd Annual SNN symposium on Neural Networks, Nijmegen, The Netherlands*, pages 366–373. Springer, 1995.
18. A. P. de Weyer. Modeling of industrial processes using natural computation. In Bert Kappen and Stan Gielen, editors, *Neural networks: artificial intelligence and industrial applications Proceedings of the 3rd Annual SNN symposium on Neural Networks, Nijmegen, The Netherlands*, pages 271–279. Springer, 1995.
19. W.J.M Epping, S.M. Oudshoff, and F.V. Abbots. Lithofacies identification from wireline logs: Bringing neural networks to application. In C.C.A.M Gielen and H.J. Kappen, editors, *Proceedings ICANN*, pages 876–881. Springer Verlag, 1993.

This article was processed using the LTEX macro package with LLNCS style