State of the Art of Neural Networks in Meteorology

Midterm Paper for course in Neural Networks

By: Bjarne Hansen

For: Dr. M. E. El-Hawary

Technical University of Nova Scotia

March 10, 1997

Bjarne K. Hansen, 1997: State of the art of neural networks in meteorology, mid-term paper for a Neural Networks course at the Technical University of Nova Scotia.

Table of Contents

Bjarne K. Hansen, 1997: State of the art of neural networks in meteorology, mid-term paper for a Neural Networks course at the Technical University of Nova Scotia.

Introduction

Neural networks have become a popular tool for solving complex problems in diverse domains over the past ten years. There are two main reasons for their popularization:

- Their demonstrable superiority at solving certain types of problems.
- Their name is alluring.

Consider how slow the English-speaking world has been to accept fuzzy logic. Pundits say that to most people the word “fuzzy” is pejorative. Poor first impressions of fuzzy logic have caused many people to not give it a second look. However, “neural network” conjures up images of machines that think as intelligently as people. Indeed, the field of neural networks has drawn bright researchers from fields as diverse as computer science, neurobiology, physics, and philosophy. Ironically, neural networks’ name has also has a detrimental effect towards their general acceptance. The notion of machines thinking like people is, to many, mysterious, unfathomable, and perhaps even disconcerting. This apprehension has been fostered by scenes from science fiction where intelligent machines, powered by neural networks, go berserk and set out to destroy their creators. This type of imagery does not encourage conventional researchers to invest their time and money in neural networks. To promote the use of neural networks, this paper will present arguments and evidence along four lines:

1. Neural networks are simple.
2. Neural networks are well suited for some types of problems, and poorly suited for others.
3. Confidence can be attached to the results of neural networks if the systems are designed carefully.
4. Superior expert systems can be built when neural networks are allied with other techniques. In particular, fuzzy-neural systems show great promise.

According to Rumelhart *et al* (1994), there are five main reason for the revived interest in neural networks that has taken place over the past ten years.

- Today's computers are much faster than those of the 1950's and 1960's.
- Everyone believes the future of computing must be in parallel computation. Neural networks are often suitable for parallel organization.
- We have learned much more about how the human brains' neurons actually work. The Holy Grail of neural networks is a machine that performs like a genius. Knowledge of the brain's actual workings lets us design better models.
- Hopfield showed that recurrent networks with symmetric weights have a point-attractor dynamics, making their behavior relatively easy to understand and analyze.
- The severe criticisms of the original perception work in the 1960's have been repudiated. Multilayer backpropagation neural networks that use gradient search methods have been shown to be universal approximators of non-linear systems.

There is a convergence of interest in neurons nowadays. In a recent edition of *La Press*, Thibaudeau (1997) covers the topic of human nerves and neurons. Advances in neurology are offering new hope to victims of motor-neural and brain disorders. But Serge Rossignol, director of the *Centre de recherche en science neurologiques* (CRSN*) de l'Université de Montréal*, cautions:

> *"Périodiquement, on crie au miracle du paralytique. Mais il faut réaliser qu'il n'y aura pas de solution simple, sous form d'une pilule, par exemple. Cela viendra plutôt de la conjunction des apports de different disciplines."*
>
> "Periodically, people proclaim a miracle for paralysis victims. But we must realize that there will not be a simple solution, as in the form of a pill, for example. It will come more from a combination of what different disciplines have to offer."

How does this relate to artificial neural networks? Ethics prevent doctors from experimenting freely with human brains and nervous systems. Artificial models serve as a model that let doctors test structures without endangering patients. Computers also have potential to serve as prostheses. In speaking of cyborgs, Ogden (1995) notes, “There are currently seventy-two ‘off-the-shelf’ body parts available.” New materials are being developed that do not provoke immune responses from graft patients. Chips are becoming continually smaller and more complex. It is reasonable to suppose that, in the not-too-distant future, electronic-nervous-implants will restore or even augment some of people’s limited neurological functions. This is still in the realm of science fiction, but most good technology stems from an inspired imagination.

Reflecting on Rossignol’s comments, we can see two points that are apropos to the topics to be covered in this paper:

- When one draws from different disciplines, one develops better techniques.
- Neurologists want to restore brain functioning; they want to restore real intelligence. In this effort, they draw from different disciplines and develop better models, such as artificial neural networks. In the same way, researchers who work with *artificial* intelligence should appreciate what “the other side” is up to. When a neurologist isolates a principle that underlies intelligence, it is totally relevant to a researcher of artificial intelligence.

Simple Nature of Neural Networks

In our context, the term "neural network" refers to a mathematical technique that maps from an input space to an output space. To do this, most neural networks have three vital components:

- Graded activation functions.
- Matrix multiplication.
- Minimization operations on inner products of vectors.

The basic workings of neural networks can be described in a few words. In practice, neural networks relate classes, patterns, or scalar values to available parameters. Input parameters consist of any perceptible signal encoded as a vector. The vector is multiplied by the weight matrices of the neural network. The product is passed to a node. If the product exceeds the node's thresholds, the node is "activated", and it conveys a signal. Nodes are interconnected in layers and networks to form neural networks. Specific patterns of input vectors elicit specific outputs from the network. Ideal values of neural network's weights and threshold values are found via relatively simple optimization techniques. After the network's weights and thresholds are optimized, training is stopped. When the trained network is presented with new input, it mimics the patterns that are present between parameters in the real world.

Zurada (1992) identifies three categories of neural networks, and describes their qualities:

> **Feedforward**: often referred to as a "back-propagation network", or simply as a "backprop"; the most commonly used form of network; undergoes supervised training; used for classification, pattern recognition, and function approximation.

Hopfield: also known as "recurrent autoassociative memory"; used primarily in pattern recognition; undergoes supervised training; after being trained, patterns are recorded in weights.

Kohonen: can be trained without supervision; useful for automatic cluster detection.

Artificial neural networks (ANN) were conceived as simple analogues of the human brain's neural structure. Reversing the analogy may help people to understand the basis of neural networks. How does someone predict that snow is imminent? They observe the world around them and learn to attach significance to certain patterns and signs. Significant signs include things such as: temperature, temperature trend, humidity, sky condition, sky condition trend, and wind direction. Certain combinations of these parameters indicate snow is likely to begin soon. Conversely, if the temperature is 20°C, one can be sure that snow is not imminent. Patterns in observations always precede certain events. People have noticed patterns in the weather for thousands of years, and become skilled at making short-term weather predictions based on their observations. This type of skilled behavior is what practitioners of neural networks intend to duplicate.

Of course, there are many refinements that enable neural network to perform more successfully. The refinements can be very detailed, complex, and fraught with *ad hoc* jargon. But it is important to keep in mind that neural networks are fundamentally simple. The add-ons that developers create are relatively minor embellishments.

How to Use a Neural Network

In principle, neural networks can solve any function, and they can do anything that digital computers can do. In practice, neural networks are mostly used for:

- Classification.
- Function approximation.
- Mapping problems where lots of training data, or cases, are available, but rules and relationships are not known or well understood.

Rumelhart *et al* (1994) offers a pointed piece advice about how to design neural networks: use "Occam's Razor." In other words, the simplest hypothesis that is consistent with the data should be chosen. Simplicity of a network can be measured in terms of: the number of weights; the number of nodes; the number of symmetries among the weights; and the number of bits per weight. Rumelhart *et al* (1994) also offers four hints for successful applications:

1. Have enough data to constrain your model sufficiently to the problem at hand.
2. Carefully design the input data. Theory-based data reduction will reduce the number of input variables. Meaningful transformations from pattern to representation space will improve performance and speed training.
3. Build known symmetries into the model. Related weights can be paired with links. Irrelevant weights can be dispensed with, or pruned.
4. Feel free to combine the network with other statistical methods. Giving clear probabilistic/Bayesian interpretation of the behavior helps to link the network with statistical methods. It helps to have the network provide output that is reasonably interpreted as probabilities. These "probabilities" can be used to determine confidence values, and can be combined by various methods with other information to build more comprehensive systems.

How Not to Use a Neural Network

Neural networks deal strictly with raw numbers. Neural network solutions are entirely determined by the contents of training data and, in some cases, by a teacher's supervision. Neural networks are not well suited for problems that involve manipulation of symbols; such problems are the forte of expert systems. When problems are expressed with symbols and symbolic manipulations, it is implied that meta-knowledge exists. Neural networks can only reflect what is inherent in raw data.

All neural networks are based on activities related to pattern recognition. Developers and users of neural networks may not be cognizant of patterns, but at some level, neural networks rely on the existence of patterns. The natural world abounds with patterns. Neural networks' unique skill at processing pattern-laden information is what is drawing researchers to the field. Scientists who deal with problems of the real world are often baffled by what to do with massive amounts of data. It is next to impossible to derive all the rules to optimally interpret the data. Neural networks may look like a tempting short cut -- a way to weed through complex, noisy data, and to isolate relationships hidden among measured variables. What we must keep in mind is that for a neural network approach to be valid, *there must be a pattern in the data*. And for the neural network results to be reliable, *the pattern must be reliable*.

When our conclusions begin with statements such as, "Assuming that patterns hold...", we must be careful. We are making three massive assumptions:

- That there is a pattern.
- That our measurements and data are sufficient to capture the pattern.
- That the system being modeled will not change significantly.

If any one of these assumptions does not hold, the neural network's output will be misleading. For instance, someone could naively assume there is a pattern between winning lottery numbers and the numbers they have seen on automobile license plates over the course of the previous week. If they wanted to, they could train a neural network to predict winning numbers based on the license plate numbers they have seen. No matter how doubtful a relationship is, a neural network's weights will converge deterministically towards a set precise values. The first person who pulls off this trick will certainly become famous (unless they wish to keep their technique a secret).

In weather prediction, there are many opportunities to misapply neural networks. Suppose we want to predict wave amplitudes in the waters off eastern Canada. For training data, we have a ten-year time series of wave measurements (predictand), and various weather parameters (predictors). A neural network can be trained to simulate the patterns that are assumed to be present in the time series. However, we know that during the past ten years, in the waters off eastern Canada, three 100-year waves have been measured. Based on statistics, waves of this amplitude are only expected to occur about once every one hundred years. Shall we base our expectations on the neural network's analysis or on the statistical analysis? The answer is not clear; therefore, it is reasonable to view both analyses with a degree of skepticism. Because weather is a chaotic and dynamic system, no particular model will always give the best predictions. Day to day, century to century, the weather changes. When features change, our models must adapt. If one asserts that all we know is all that can exist or can be known, one risks of being accused of paradigmatic solipsism.

Rumelhart *et al* (1994) describe an efficient way to test the validity of neural network results and to avoid the problem of overfitting: use a scheme of cross-validation. Divide the training data into three parts: a training set; a set to test the generalization performance during testing; and a cross-validation set to test the performance of the systems after training. Through cycles of training, or epochs, the system trains on the first set, and tests its skill at generalization on the second set. When generalization begins to deteriorate, the systems is beginning to overfit to the first set, and training is stopped.

Finally, the performance of the system is tested with the reserved cross-validation test. This method is simple and often leads to good results.

Neural networks usually rely on huge amounts of representative data to be trained. If one lacks sufficient data, the neural network approach is inappropriate. And even if one does have a large stockpile of data, that alone is not enough. The data needs to be judiciously prepared before it is submitted to a neural network for use as training data. Simply entering raw data into a neural network in a willy-nilly way is bound to lead to disappointing results. The most challenging parts of neural network design are often the collection and the preparation of suitable training data.

Confidence in Black Box Solutions

Neural networks are excellent at recognizing patterns; it is what they are designed to do. Patterns are stored in the weights of the networks. The weights are derived by an iterative mathematical technique, not by any method resembling classical logic. It is understandable when people inherently mistrust neural network solutions; the weight optimization procedure is so unlike how intelligent people think. The inner workings of neural networks are unseeable for most people; hence, neural networks are often referred to with the pejorative nickname, "black box."

The absence of reasoning in neural networks appears to many people to be their main weakness: neural networks' results are not based on infallible logic. Worse yet, their results are unstable; tiny variations in training data can lead to widely differing solutions. Neural networks are comparable to statistics: there is very little logic in statistics; and tiny variations in the sample can have profound impacts on the conclusions. Yet, we regularly base important decisions on statistical analysis. Statistics leads us to make statements like, "Based on our observations, and assuming that patterns hold, we can expect…" The style of reasoning used with statistics is virtually identical to that used with neural networks. The operations that are used in statistics run parallel to those used in neural networks. Neural networks should seem no more esoteric than linear regression. In fact, neural networks are often referred to as "nonlinear regression." Schurmann (1996) goes so far as to say that neural networks and statistics are "inextricably fused." He has been working for twenty years in the fields of statistics and pattern recognition. He argues persuasively that superior pattern recognition systems integrate neural networks and statistics, and thereby acquire the strengths of both techniques.

The lack of agreed-upon confidence measures for neural network solutions also is viewed by many as a major weakness. Several research articles reviewed in this paper grapple with issue of confidence. Nobody claims to have the ultimate definition of "confidence in neural network output." Besides, only an oracle would have the nerve make such a statement. The current state of neural network development is reminiscent of the state of statistics in the early part of this century. People often overlook the history of statistics. It took decades of research and development before statistics gained its current image of respectability. In recent years, there has been a boom of research and development in the field of neural networks. If this pattern holds, it is likely that neural network text books will become as familiar as statistical texts. And it probably will not take decades for this to happen. Neural networks today have two advantages that statistics lacked in its formative years: computers have accelerated numerical research; and the field of statistics itself provides a firm foundation for the development of neural networks.

Meteorological Applications of Neural Networks

Reports of neural networks began to appear rather suddenly in the meteorological literature three years ago. Interest in the technique has grown steadily since then. It now seems that it was inevitable that meteorologists would adopt neural networks as a standard technique. Consider what Conway (1989) wrote eight years ago in his paper about expert systems and weather forecasting:

> "An important observational source, particularly for short-range forecasting, is remotely sensed imagery. Much is often made of the difficulties of automatically recognizing patterns in images … One of the important functions of an expert system will be to use other meteorological knowledge and evidence to focus and constrain the pattern-recognition task, and avoid getting bogged down in huge amounts of low-level image-processing."

Conway identifies the main weaknesses of meteorological expert systems: an inability to efficiently recognize patterns. Conway argues that pattern recognition is an essential component of weather forecasting systems; therefore, he proposes that search algorithms be applied to pattern-containing databases. Developing such algorithms is exceedingly difficulty because the pattern space is huge. In 1989, neural networks were almost unknown to meteorologists. Now, it is obvious that neural networks are ideally suited to the pattern recognition problem described by Conway.

Highlights of twelve neural network-based meteorological articles are summarized here. The list below is not comprehensive; the number of meteorological applications of neural networks is increasing faster than one can easily keep track of. Rather, the list should be seen as evidence that neural networks are well suited for a wide range of pattern-recognition-related meteorological problems. The articles are rated subjectively, ranging from good (★★) to great (★★★★★). In some cases, the articles prompted ideas (➔) or criticism (☹).

Minimum Temperature Forecasting ★★★

Author:	Schizas *et al* (1991)
Type of network:	backprop
Topology:	?
Input data:	141 successive days of: total cloud cover; eastward wind component; northward wind component; visibility; present weather condition; atmospheric pressure; dry-bulb temperature; wet bulb-temperature; low cloud cover; day length; and observed minimum of previous night
Results:	with all input: ≈18% accurate ± 0.5°C
	≈68% accurate ± 3.0°C
	with only: present weather condition; dry-bulb temperature; wet bulb-temperature; observed minimum of previous night
	≈21% accurate ± 0.5°C
	≈66% accurate ± 3.0°C
Key points:	12 combinations of input parameters tested
	four parameters alone offer good results
	ANN is site specific

Fog Forecasting ★★★

Author:	Naranjo-Diaz *et al* (1995)
Type of network:	Kohonen - Learning Vector Quantization
Topology:	-
Input data:	1600 UTC temperature and wind speed previous day; and occurrence of fog on previous day (persistence)
Results:	with large data sets, neural net beats discriminant analysis
Key points:	output is YES or NO prediction of morning fog in a region best to train sites with different fog climatologies separately ➔ input site as a coded parameter, network may learn climatology

Rainfall Estimation from Satellite Imagery ★★★★

Author:	McCullagh *et al* (1995)
Type of network:	backprop
Topology:	$4 \rightarrow 4 \rightarrow 1$, $54 \rightarrow 16 \rightarrow 1$, and $58 \rightarrow 20 \rightarrow 1$
Input data:	satellite data (cloud top temperatures and brightness); NWP data; and combination of satellite and NWP
Results:	neural net beats both GOES Precipitation Index Technique and Australian Regional Assimilation Prognosis
Key points:	determination is RAIN or NO RAIN in southeast Tasmania satellite data alone can be used for climate monitoring satellite-derived values relate to values measured on the ground different systems have highest skill in different situations neural net good for cumiloform, poor for stratoform and cirrus intend to refine technique for quantitative precipitation forecast will investigate fuzzy classification technique ☹ training set very small, cannot contain full range of patterns

Snowfall and Rainfall Forecasting from Weather Radar Images ★★★★★

Author:	Ochiai *et al* (1995)
Type of network:	backprop
Topology:	$25 \rightarrow 3 \rightarrow 1$ with 57, 600 parallel networks
Input data:	time series of weather radar
Results:	neural network beats persistence and cross correlation methods
Key points:	parallel organization makes it feasible to: train "on-the-fly" learn current dynamics, and accommodate local geography "kick-out" algorithm lessens problem of weight oscillation

Lightning Forecasting ★★★

Author:	Frankel *et al* (1991)
Type of network:	backprop
Topology:	$106 \rightarrow ? \rightarrow 64$
Input data:	wind field, electric field, satellite cloud-top altitudes, and "others"
Results:	neural network beats prior state-of-the-art method
Key points:	development was prompted by a lightning event that caused NASA to lose a $160 million satellite 😐 training set only two days, cannot contain full range of patterns

Tornado Prediction ★★★★★

Author:	Marzban *et al* (1996)
Type of network:	backprop
Topology:	start with two hidden layers, then refine with pruning
Input data:	23 variables relevant to tornadoes derived from Doppler radar
Results:	neural network beats: rule-based algorithm, discriminant analysis, and logistic regression
Key points:	excellent treatment of statistics prediction is YES or NO that mesocyclone will spawn tornado show how pseudo-probabilities can be extracted from network confidence in neural network can be based on pseudo-probabilities ➔ fuzzification of input and defuzzification of output may allow us to configure an algorithm to associate confidence in results e.g. YES = 0.2 & NO = 0.8 $\Rightarrow$ probability = 30%

Hurricane Tracking ★★

Author:	Johnson, G.P. *et al* (1995)
Type of network:	backprop
Topology:	$8 \rightarrow 1 \rightarrow 8$
Input data:	time series of hurricane data: time, latitude, longitude, speed, direction, maximum wind, minimum pressure, cyclone stage
Results:	comparable to other National Hurricane methods, but not superior
Key points:	training is done storm-by-storm ☹ training set consists of only hurricanes of one season, dynamical nature of hurricanes implies that two are rarely similar; furthermore, there is much variation from one season to another

Cloud Recognition with Ground Sensors ★★

Author:	Aviolat *et al* (1996)
Type of network:	backprop
Topology:	? → ? → ? automatic construction by constructive learning and pruning techniques
Input data:	standard set of weather data from auto-stations, pyrgeometer, and human cloud estimates (truth)
Results:	estimates total cloud with accuracy ≈ 80-90%
Key points:	system being built is Swiss equivalent of AWOS object is to have robot code METAR messages present view of unified heuristic-neural expert systems ANN does not differentiate between layers ☹ system is trained with mixture of low, mid, and high cloud yet input selection based on assumption that "evolution of meteorological phenomena are slow"; this is not true of low cloud, and low cloud is most significant for pilots in the Alps

Cloud Classification with Remote Sensing ★★★★

Author:	Bankert (1994), Bankert *et al* (1996)
Type of network:	backprop
Topology:	15 → ? → ? → 1
Input data:	200 various features measured by NOAA polar orbiting satellite; optimal set of 15 is selected; truth is set of subjectively labelled images
Results:	rated against subjective method, 77-98% accurate, depending upon stringency of tests
Key points:	classify cloud genus directly from satellite data training done by hold-one-out method second paper employs greedy algorithm for feature selection trade-off between accuracy and precision: when output precision is broadened from one-of-ten to one-in-five, accuracy increases ➔ fuzzification of input and defuzzification of output may allow us to configure an algorithm to associate confidence in results [*via email, author suggested this to Bankert, and Bankert agreed*]

Identification of Sea Ice Coverage and Movement with Remote Sensing ★★★

Author:	Rau *et al* (1994)
Type of network:	backprop to classify ice field, Hopfield to store classified images and assess change over time
Topology:	4 → 200 → 100 → 8
Input data:	NOAA polar orbiting satellite images, wind
Results:	inconclusive, research in ongoing
Key points:	classify 8 types of ice detect ice movement by comparing successive classified images

El Nino Prediction ★★★

Author:	Derr *et al* (1994)
Type of network:	backpropagation feedforward
Topology:	12 → 30 → 1
Input data:	COADS database, 420 successive months of data: Southern Oscillation Index, sea surface temperature, east-west component of wind
Results:	neural net beats persistence for all but shortest lead times
Key points:	object is to predict onset of El Nino events (Pacific heat anomaly) deals with forecasting of climate El Nino → global weather patterns altered

Ensemble Forecast Classification ★★★★

Author:	Eckert *et al* (1996)
Type of network:	Kohonen
Topology:	-
Input data:	ensemble of 32 ECMWF long range 500 hPa predictions
Results:	inconclusive, research ongoing
Key points:	object to detect clusters (similar weather) among 32 ensembles wrap-around strategy used to avoid spurious aggregations along edges and corners of neural chart

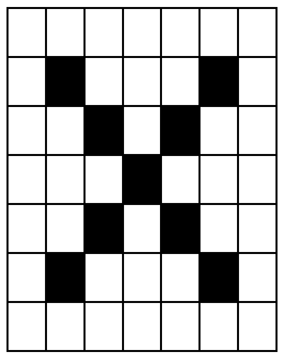
Fuzzy-Neural Hybrid Systems

Srinivasan *et al* (1994) develop a fuzzy-neural system for forecasting electrical load requirements. The fuzzy-neural network's results are significantly more accurate than any of the other systems used to predict load. The competition includes: neural network (alone); multiple regression; and auto-regressive moving average. Srinivasan *et al* attribute the success of their system to seven of the system's characteristics:

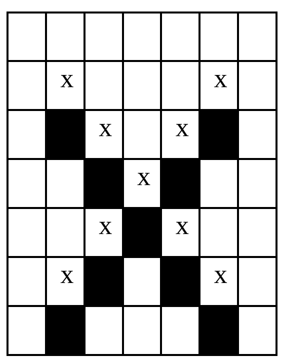
1. Amalgamation of different pieces of knowledge is possible by application of fuzzy rules.
2. A large-scale knowledge base can be effectively handled and reduced by a fuzzy front-end processor [the fuzzy add-on to the ANN], making learning easy and fast.
3. Non-precise and context-dependent knowledge is represented using fuzzy logic.
4. Recognition and learning from noisy data is possible.
5. The neural network processes the data and knowledge by combining these two, constructed inductively and independently.
6. The technique is robust in that only some rules in the fuzzy knowledge base are required to be updated with changing input conditions, avoiding the need to retrain the neural network.
7. Neural network learning is able to progressively generalize knowledge from many instances. Therefore, if the knowledge of the inference is constructed using the fuzzy front-end processor, the inference becomes smarter every time new type of data is presented to the network.

Recognizing Shifted and Distorted Patterns

Kwan *et al* (1994) describe how pattern recognition becomes more accurate when fuzzy methods are combined with neural networks. Their work deals with optical character recognition. They show how fuzzy logic lets the network contend with shifted patterns. A simplified version of their examples shows how fuzzy logic allow shifted patterns to be recovered. Suppose the letter “x” is stored in a pixel-network in this manner:

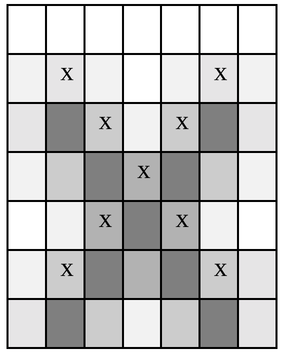


Then suppose the “x” pattern is shifted.

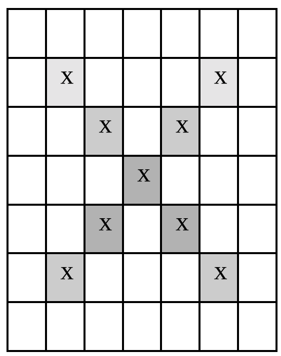


Pixel-to-pixel, the second image is very unlike the first. None of the pixels that correspond to prototypical “x” positions, marked by “x”, appear black. According to a crisp interpretation of the pixels, the image of “x” is lost.

Notice what happens when the second image is fuzzified.



When we focus our attention on the pixels that signify a prototypical “x”,



we see that all the *x-to-x* pixels are active again, albeit, to a degree less than 1.0.

The pattern recognition problems that Kwan *et al* solves are of a realistic size. To evaluate their system, Kwan *et al* compare their system’s output with a Hamming network’s output. A Hamming network classifies an input pattern by measuring the Hamming distances between the input pattern and all the training patterns. The fuzzy-neural network is significantly better that the Hamming network at recognizing shifted and distorted patterns.

Ozone Prediction

Burrows *et al* (1997) describes an ozone prediction application based on the Neuro-Fuzzy Inference System (NFIS). NFIS improves the performance of a neural network by using a method termed "fuzzy clustering" to reduce the dimensionality of the input data. Thousands of widely distributed data points are available for training. Before the data are input to the neural network, they are mapped into a manageable set of clusters. The cluster selection procedure uses fuzzy methods to progressively determine the optimal cluster centres. After clusters have been identified, one can develop sets of fuzzy rules, each set of rules pertaining to one set of clusters. Via email, Borrows elaborates on the cluster identification process:

> "The basis for choosing a cluster point is a 'potential' which is calculated as a function of distance from a cluster center and a radius of influence. The process is iterative. After a cluster center is determined, the potential of all other points is adjusted by subtracting an amount dependent on their distance from other cluster centers already found. The point with the 'highest' adjusted potential is a candidate for the next cluster center, not the point that is most distant from points already chosen as cluster centres. The cluster centers are vector centers, that is, they include all components, and distances, calculated as vector distances."

Discussion

How should we select an ideal strategy to solve a problem? Let us draw an example from nature: how should frogs solve the problem of obtaining food? Hypothetically, they could use two strategies:

- Frogs could "reason" about how to obtain food. They could methodically search the space around them, until they achieved their objective.
- Or they could automatically respond to certain patterns: "When a black dot moves into view, grab it."

Natural selection has chosen the second strategy. If frogs had to "reason" about how to respond to black dots flying into view, by the time they decided to grab the fly, the fly would be gone. From natural selection's point of view, logic is an unusual strategy.

The articles reviewed in this paper are all concerned with the selection of optimal strategies for solving meteorological problems. All the papers have three things in common:

- They evaluate neural networks.
- They determine that the skill neural networks is either comparable or superior to other methods.
- They urge that continued development of neural networks be pursued.

In his paper describing a fuzzy expert system for meteorology, Hansen (1997) concludes:

> "Meteorological information and knowledge are often uncertain, ambiguous, or vaguely defined. Operational meteorology is therefore treated as a fuzzy environment. An argument is made for the applicability of methods based on fuzzy logic for the optimal solution of problems related to the evaluation of

> meteorological data and forecasts. An expert system has been designed which uses fuzzy methods to interpret meteorological data. The system efficiently identifies significant information contained within huge amounts of data. Forecasters using the system can more consistently and easily monitor the accuracy of weather forecasts. Systems such as [the fuzzy expert system] described here are bound to become more common as time goes on."

According to Zurada (1992), a hallmark of first-rate neural networks is a well-designed procedure for mapping input vectors from a pattern space into a representation space. Such a procedure has three effects:

- Significant features in the pattern space are highlighted.
- Insignificant signals, such as noise, are suppressed.
- Dimensionality is reduced.

These are exactly the types of procedures for which fuzzy logic is intended. The case for designing systems that fuse fuzzy logic and neural networks is compelling.

Neural networks are commonly treated as a distinct and isolated branch of artificial intelligence. People who attempt to solve problems using artificial intelligence often rely on only one approach; but from one case to another, the best approach will vary. In many cases, one technique is sufficient; but in many other cases, a combination of techniques will yield best results.

Conclusion

This is a pivotal decade in the evolution of meteorological expert systems. AI gives us a diverse array of techniques for solving problems; each technique has its characteristic strengths and weaknesses. Neural networks and fuzzy logic are two such techniques. For problems that belong to clearly defined categories, one technique is often the best. However, realistic problems seldom fit into clear categories; they often break down into sets of unique sub-problems. In these cases, the optimal technique for solving the combination of problems relies on a combination of AI techniques. Based on the arguments and examples given in this paper, we believe that fuzzy-neural expert systems are an important type of hybrid-AI technique. The unique combination of strengths that fuzzy-neural systems possess help them to outperform many other type of expert systems.

References

1. Aviolat, F., Cattani, D., and Cornu, T. - "Recognition of Meteorological Situations with Neural Networks", Proceedings of the Third Biennial Joint Conference on Engineering Systems Design & Analysis, ESDA '96, Montpellier, July, 1996

2. Bankert, R.L. - "Cloud Classification of AVHRR Imagery in Maritime Regions Using a Probabilistic Neural Network", *Journal of Applied Meteorology*, Vol. 33, August 1994, pp. 909-918

3. Bankert, R.L., and Aha D.W. - "Improvement to a Neural Network Cloud Classifier", *Journal of Applied Meteorology*, Vol. 35, November 1996, pp. 2036-2039

4. Burrows, W.R., and Pudykiewicz, J. - "CANFIS: A Non-Linear Regression Procedure to Produce Statistical Air-Quality Forecast Models", to be presented at the Air & Waste Management Association's 90th Annual Meeting & Exhibition, June 8-13, 1997, Toronto, Canada, (manuscript via personal communication)

5. Conway, B.J. - "Expert Systems and Weather Forecasting", *Meteorological Magazine*, Vol. 118, No. 1399, February 1989, pp. 23-30

6. Derr, V.E., and Slutz, R.J. - "Prediction of El Nino Events in the Pacific by Means of Neural Networks", *AI Applications*, Vol. 8, No. 2, 1994, pp. 51-63

7. Eckert, P., Cattani, D., and Ambuhl, J. - "Classification of Ensemble Forecasts by Means of an Artificial Neural Network", *Meteorological Applications*, No. 3, 1996, pp. 169-178

8. Frankel, D., Schiller, I., Draper, J.S., and Barnes, A.A. Jr. - "Use of Neural Networks to Predict Lightning at Kennedy Space Center", International Joint Conference on Neural Networks, 1991 (in Seattle, Wa.), IEEE, New York, New York, pp. 319-324

9. Hansen, B. - "SIGMAR: A Fuzzy Expert System for Critiquing Marine Forecasts", *AI Applications*, Vol. 11, No. 1, 1997 (in press)

10. Kwan, H.K., and Cai, Y. - "A Fuzzy Neural Network and its Application to Pattern Recognition" - IEEE Transactions on Fuzzy Systems, Vol. 2, No. 3, August 1994, pp. 185-192

11. Johnson, G.P., and Lin, F.C. - "Hurricane Tracking by Backpropagation Neural Network", Proceedings of the 1995 IEE International Conference on Neural Networks, Nov. 27-Dec. 1, 1995 (in Perth, Australia), IEEE, New York, New York, U.S.A., pp. 1103-1106

12. Naranjo-Diaz, L.R., and Alfonso, A.P. - "Fog Forecasting in Cuba. Neural Networks versus Discriminant Analysis", *Meteorological Applications*, No. 1, 1995, pp. 31-33

13. Marzban, C., and Stumpf, G.C. - "A Neural Network for Tornado Prediction Based on Doppler Radar-Derived Attributes", *Journal of Applied Meteorology*, Vol. 35, May 1996, pp. 617-626

14. McCullagh, J., Bluff, K., and Ebert, E. - “A Neural Network for Rainfall Estimation”, Conference Paper, Proceedings of the 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, IEEE Computing Society Press, Los Alamitos, Ca. U.S.A., pp. 389-392

15. Ochiai, K., Suzuki, H., Shinozawa, K., Fujii, M., Sonehara, N. - “Snowfall and Rainfall Forecasting from Weather Radar Images with Artificial Neural Networks”, Proceedings of the 1995 IEE International Conference on Neural Networks, Nov. 27-Dec. 1, 1995 (in Perth, Australia), IEEE, New York, New York, U.S.A., pp. 1182-1187

16. Ogden, F. - Navigating in Cyberspace, MacFarlane, Walter & Ross, Toronto, 1995, pg. 143

17. Rau, Y.-C., Cosimo, J.C., and Lure, F.Y.M. - “Application of Neural Networks for Identification of Sea Ice Coverage and Movements from Satellite Imagery”, Proceedings of the International Geoscience and Remote Sensing Symposium. Surface and Atmospheric Sensing: Technologies, Data Analysis and Interpretation (August 8-12, 1994 in Pasedena, Ca., USA), IEEE, New York, New York, USA, pp.1407, 1409

18. Rumelhart, D.E., Widrow, B., and Lehr, M.A. - “The Basic Ideas in Neural Networks”, *Communications of the Association for Computing Machinery*, Vol. 37, No. 3, March 1994, pp. 87-92

19. Thibaudeau, C. - “Les Neurones: Repousseront-ils Un Jour” (Neurons: Regrow Them One Day), *La Press* (Montreal newspaper), March 9, 1997, pp. C1-C2

20. Schizas, C.N., Michaeldes, S., Pattichis, C.S, and Livesay, R.R. - “Artificial Neural Networks in Forecasting Minimum Temperature”, Second International Conference on Artificial Neural Networks, 18-20, Nov. 1991, IEEE, London, U.K., pp. 112-114

21. Schurmann, J. Pattern Classification - A Unified View of Statistical and Neural Approaches, John Wiley & Sons, New York, New York, U.S.A., 1996

22. Srinivasan, D., Liew, A.C., and Chang, C.S. - “Forecasting Daily Load Curves Using a Hybrid Fuzzy-Neural Approach”, IEE Proc.-Gener. Transm. Distrib., Vol. 141, No. 6, November 1994, pp. 561-567

23. Zurada, J.M. Introduction to Artificial Neural Systems, West Publishing Company, St. Paul, Mn., U.S.A., 1992