

Simulation Studies of Data Classification by Artificial Neural Networks: Potential Applications in Medical Imaging and Decision Making

Yuzheng Wu, Kunio Doi, Charles E. Metz, Naoki Asada, and Maryellen L. Giger

Artificial neural networks are being investigated in the field of medical imaging as a means to facilitate pattern recognition and patient classification. In the work reported here, the effects of internal structure and the nature of input data on the performance of neural networks were investigated systematically using computer-simulated data. Network performance was evaluated quantitatively by means of receiver operating characteristic analysis and compared with the performance of an ideal statistical decision maker. We found that the relatively simple neural networks investigated in this study can perform at the level of an ideal decision maker. These simple networks were also found to learn accurately even when the training data are extremely unbalanced with respect to the prevalence of actually positive cases and to differentiate input data patterns by recognizing their unique characteristics.

Copyright © 1993 by W.B. Saunders Company

KEY WORDS: neural network, receiver operating characteristic (ROC) analysis, classification, ideal decision-maker, signal-to-noise ratio (SNR), simulation

ARTIFICIAL NEURAL networks provide a nonalgorithmic approach to information processing. This approach can solve problems not by use of a prespecified algorithm, but rather by "learning" from examples presented repeatedly. An artificial neural network consists of a number of processing units that are layered and interconnected much like neurons in the human brain. Neural networks have been shown to be a useful tool for pattern recognition and classification in fields in which conventional algorithmic approaches and rule-based expert systems may not be successful.^{1,2}

In recent years, neural networks have been applied in the field of medical imaging³⁻¹² to facilitate pattern recognition and patient classification in a variety of different situations. In chest radiography, for example, the differential diagnosis of interstitial lung diseases is an important but difficult task for radiologists. Accurate interpretation of chest films in this task involves the merging of diverse radiographic findings and clinical data to classify patients. Because the mapping between patterns of image features and states of health and disease may be

ill-defined or unknown, it is difficult to design and implement conventional algorithms to assist radiologists in performing this task. Hence, neural networks have been used in an attempt to differentiate among patterns of input data that correspond to different interstitial diseases.^{6,7} Other applications of neural networks in medical imaging involve x-ray spectral reconstruction from measured spectra,⁸ interpretation of neonatal chest radiographs,⁹ magnetic resonance imaging tissue classification,¹⁰ detection of microcalcifications from digital mammograms,¹¹ and classification of mammographically evident lesions as benign or malignant.¹²

We conducted a computer-simulation study to systematically investigate the performance of neural networks with different kinds of input data and to determine the effect of the networks' structure on their performance. Values of random variables sampled from multivariate Gaussian distributions were used to represent clinical input data. The performance of the neural networks was evaluated quantitatively by means of receiver operating characteristic (ROC) analysis.¹³ We attempted to optimize the performance of the networks by choosing internal parameters to obtain the highest ROC curves for given input data patterns.

METHOD

Three-layer, feed-forward neural networks were used in this study. The general structure of such neural networks is shown in Fig 1. These networks have an input layer, a hidden layer, and an output layer. Input data and desired

From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, MC2026, The University of Chicago, Chicago, IL.

Supported by United States Public Health Service grants nos. CA24806 and CA48985 and by an American Cancer Society Junior Faculty Award (JFRA-212).

Address reprint requests to Yuzheng Wu, MSc, Department of Radiology, MC2026, The University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637.

*Copyright © 1993 by W.B. Saunders Company
0897-1889/93/0602-0002\$3.00/0*

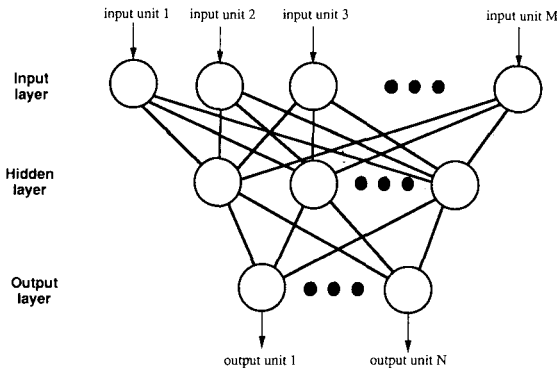


Fig 1. Structure of a typical three-layer, feed-forward artificial neural network.

output values are provided to the input layer and to the output layer, respectively. All of the input values are normalized to the range from 0 to 1, and the output values also range from 0 to 1. The hidden layer is the key component in which learning is considered to take place. Each unit in the hidden and output layers receives, as its input, weighted outputs from all of the units in the previous layer. Each such input is added to an offset value and subsequently converted to the unit's output by a sigmoidal function. The weights and offset values are determined by a training process. The numbers of units used in the input and output layers are determined by the number of input data components used and by the nature of the desired output, respectively. The number of hidden units is generally selected empirically for each particular application.

In the process of training a neural network, the output values obtained from the provided input data are compared with the desired outputs. Error is defined as the squared magnitude of the difference between the desired and calculated output vectors, summed over all training examples. Internal parameters of the neural network are adjusted repeatedly until a predetermined number of iterations has been completed or the error becomes less than a predetermined value. We used an error back-propagation algorithm in the training process and the generalized delta rule in back-propagating error. Details of back-propagation and the generalized delta rule can be found elsewhere.¹⁴ In the process of testing a neural network, outputs are calculated from a given set of input data by

using the internal parameters that were determined during the training process. ROC analysis was used to evaluate the performance of the neural networks.

In our neural network applications, each input unit is provided with a numerical value that represents one radiographic finding or clinical feature. Typical input data in mammography, for example, may include the shape and size of a mass lesion, degree of spiculations of mass, uniformity of density of mass, and distribution of microcalcifications. The values of these input features can be modeled by joint outcomes from a multivariate gaussian curve distribution, with each feature having different means and, in general, different standard deviations (SDs) for each state of "truth." A simple network with two input units and one output unit is shown in Fig 2. With single output units, the training process seeks to produce output values of 0 and 1 from input data associated with actually negative and actually positive cases, respectively.

In situations that involved two states of truth, we evaluated the performance of each neural network by comparing it with that of an ideal decision maker (see Appendix). For each feature, the relationship between the pair of Gaussian input value distributions that corresponds to the two states of truth can be characterized by a signal-to-noise ratio (SNR) defined such that "signal" ($\Delta\mu$) is the difference between the means of the distributions for the two states of truth, whereas "noise" (σ) is the root-mean-square SD of those two distributions. In this

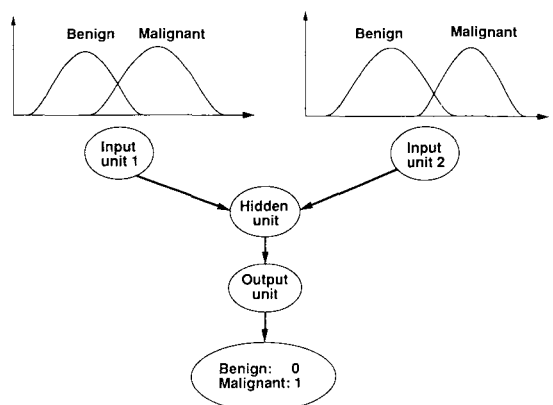


Fig 2. A neural network with two input units, and distributions of the input data used for training and testing.

report, the SDs of the Gaussian distributions for actually negative and actually positive cases were equal. If the data for N input units are statistically independent and are sampled from Gaussian distributions with the same SNR, the area (A_z) under the ROC curve that would be achieved by an ideal statistical decision maker in distinguishing between the two states of truth is related to the SNR of a single input unit by the following equation:

$$A_z = \Phi \left(\sqrt{\frac{N}{2}} \frac{\Delta\mu}{\sigma} \right) = \Phi \left(\sqrt{\frac{N}{2}} \text{SNR} \right), \quad (1)$$

where Φ represents the cumulative standard normal distribution function (see Appendix). ROC curves achieved by the neural networks were estimated by calculating true-positive fraction (TPF) and false-positive fraction (FPF) as a function of the threshold setting that distinguished nominally positive from nominally negative values at the output unit.^{13,15,16}

In each simulation study a training data sample contained a preselected number of independent training data sets. Each data set represented the input data that might be obtained from a particular patient and consisted of input values that were sampled independently from Gaussian distributions with mean values that were determined by the "truth" of the simulated case. The training data sets in a given sample were entered into the neural network repeatedly in randomly varied sequences. The number of sequences used during training is called the number of iterations. Testing data sets were sampled independently in the same way from the same distributions. Unless otherwise specified, an equal number of actually negative (benign) and actually positive (malignant) cases was included in each training or testing data sample.

RESULTS

Neural Network With Two Input Units

The simplest neural network consists of two input units, one hidden unit, and one output unit. Using a neural network with this simple structure, we investigated the relationship between statistical properties of the input data and the output obtained from the neural network. Effects of the number of iterations on the

performance of the neural network were also studied. The understanding of neural networks gained from characteristics of this simple model may serve as a guide in practical applications of neural networks with more complicated structure.

Effect of training parameters. The simple neural network with two input units was first trained by different numbers of data sets, with 200 iterations used for each, and then tested with 1,000 data sets obtained independently from the same distributions. Better performance was expected with larger training data sets, which reveal the statistical relationship between input patterns and "truth" more reliably. Figure 3A shows the relationship between the performance of the neural network and the number of input data sets that were used for its training. The network's ROC curve improved steadily as the number of input data sets increased from two to 100, with no appreciable improvement thereafter. Hence, in this simple situation, 100 data sets seem to provide sufficient training.

The effects of sampling variation in the training data, random sequencing of the training data sets, and the number of iterations were also studied. When tested with a single sample of 1,000 data sets, the performance of the neural network remained unchanged when it was trained with four different independent samples of 100 data sets. Therefore, in this simple situation the network does not appear to be sensitive to sampling variation in the training data, again indicating that 100 training cases are adequate. The effect of the sequence in which input data sets were presented during training was studied by training the neural network with various sequences of the 100 data sets in a single input sample. We found that the performance of the network was not affected. Finally, the neural network was trained by a single sample of 100 data sets and increasing numbers of iterations. We found that this simple network was trained sufficiently with only 10 iterations.

Effect of testing parameters. After being trained with 100 data sets and 200 iterations, the neural network was tested by assessing its performance repeatedly with different numbers of testing data sets. When the number of testing data sets in each simulated experiment was increased to 1,000 or more, statistical variation

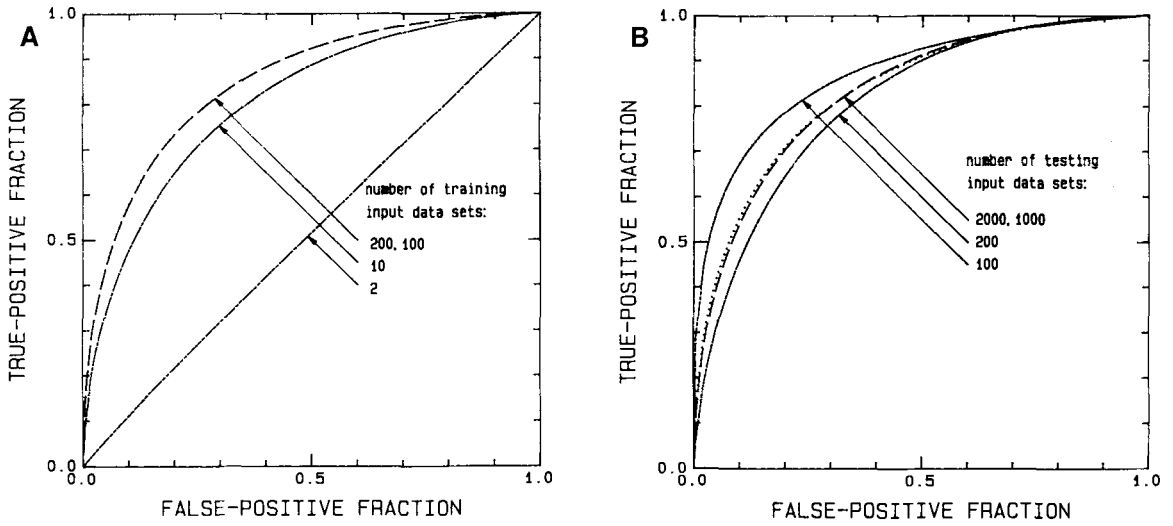


Fig 3. Effect of number of input data sets on the performance of neural networks of two input units. (A) Effect of number of training data sets: SNR [6ed] = 1, 200 iterations, and 1,000 testing data sets. (B) Effect of number of testing data sets: SNR = 1, 200 iterations, and 1,000 training data sets.

among the ROC curves and A_2 values obtained from individual simulated experiments was small. With 100 or 200 testing data sets in each simulated experiment, the estimated ROC curves and corresponding A_2 values varied as expected,^{15,16} but the average values did not change substantially. Typical ROC curves obtained from individual simulated experiments with different numbers of testing data sets are shown in Fig 3A.

Effect of unbalanced prevalence. In practical situations, different numbers of actually negative and actually positive clinical cases with established clinical truth may be available for training a neural network. Therefore, we examined the effect of unbalanced training data on the performance of the neural network by including only one to 10 actually positive (malignant) cases in a total of 1,000 input data sets used for training. However, testing data sets

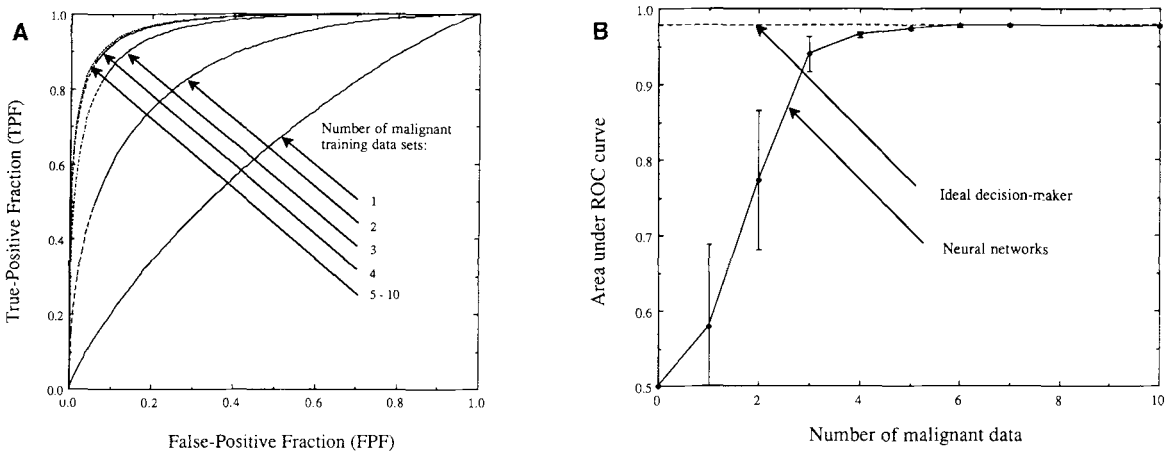


Fig 4. Effect of unbalanced prevalence in the input data sets on the performance of a neural network with two input units. SNR = 2, 200 iterations, 1,000 training data sets, and 1,000 testing data sets. (A) ROC curves achieved by the neural network with 1 to 10 malignant training cases, averaged over 14 replications of the experiment. (B) Average A_2 values achieved by the neural network with one to 10 malignant training cases (points) and the performance of an ideal decision maker (dashed line). Error bars indicate ± 1 standard deviation of the mean.

remained evenly divided between benign and malignant cases. The averaged ROC curves shown in Fig 4A indicate that the performance of the neural network improved gradually as the number of malignant training cases increased from one to 10. The corresponding average A_z values shown in Fig 4B indicate that the neural network performed at the theoretical upper limit of an ideal decision maker even when as few as five malignant cases were included among the 1,000 training data sets. This suggests that, in such simple situations, the neural network is very efficient in learning from even a small number of positive cases. However, variation among small samples of training cases can cause network performance to depend strongly on the particular training data sample that is used, and that a substantially larger number of training data sets may be required in situations that involve more than two input units, more complicated input data distributions, and/or lower SNR.

Relationship between A_z and SNR. Figure 5 shows two different pairs of bivariate input data distributions. In Fig 5A, the mean input values are larger for malignant cases than for benign cases at both input units. In Fig 5B, however, the mean input value for malignant cases is

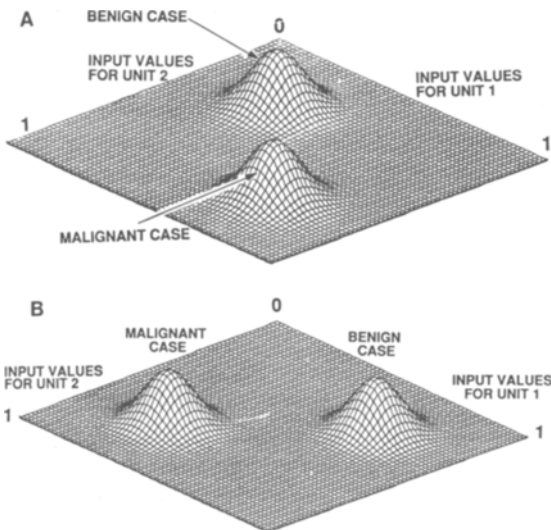


Fig 5. Bivariate binormal distributions of input data for a neural network with two input units. (A) Mean input values for malignant cases are larger than those for benign cases at both input units. (B) Mean input value for malignant cases is larger than that for benign cases at input unit 2 but smaller at input unit 1.

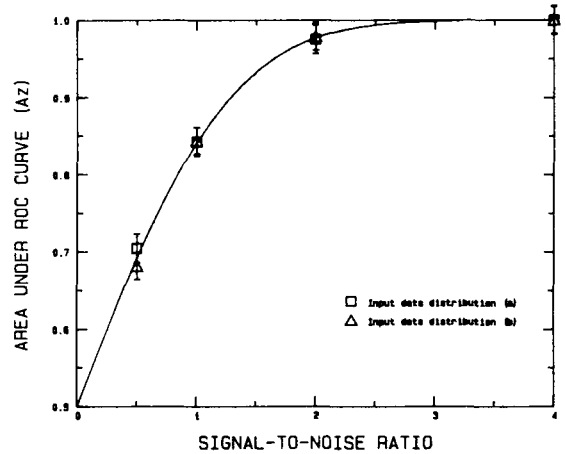


Fig 6. Dependence of A_z on the SNR at each of two input units for the distributions shown in Fig 5. Results obtained by the neural network (200 iterations, 100 training data sets, 1,000 testing data sets) are shown as points; curve indicates the theoretical performance of an ideal decision maker in both situations.

larger than that of benign cases at one input unit and smaller at the other. Our simple neural network was trained by 100 data sets and tested by another 1,000 data sets in this situation. The corresponding value of A_z for an ideal decision maker was calculated theoretically (see Appendix) as a function of SNR in the input data distributions. In Fig 6, it is apparent that the A_z values that were achieved by the neural networks correspond with those obtained from theoretical predictions for both pairs of distributions.

Neural Networks With Multiple Input Units

Equation 1 indicates that, for a given SNR at each input unit, the ideal decision maker obtains larger values of A_z if the number of independent input units is increased. This prediction is confirmed by Fig 7 which shows the performance of neural networks with different numbers of independent input units. As noted in the figure legend, the numbers of hidden units used varied correspondingly with the number of input units chosen to achieve optimal performance. The networks seem to perform at levels very close to those of an ideal decision maker in each situation. However, as noted in the figure legend, the neural networks needed more training data sets to achieve optimal performance when the number of input units was increased.

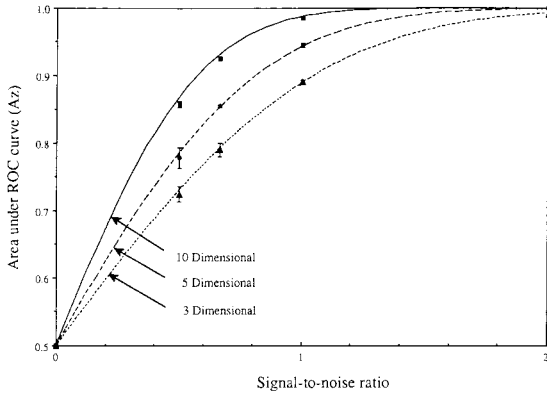


Fig 7. Performance achieved by neural networks with three to 10 input units (points) and the performance of an ideal decision maker (curves) as a function of the SNR at each input unit. As the number of input units increased, more training sets (T), more iterations (I) and more hidden units (H) were needed to approach maximum performance, as follows: for 3 input units, $T = 200$, $I = 200$, $H = 1$; for five input units, $T = 5,000$, $I = 1,000$, $H = 3$; for 10 input units, $T = 10,000$, $I = 1,000$, $H = 5$.

Neural Networks With Multiple Output Units

In many clinical applications of neural networks, the desired output may not be simply binary (eg, either benign or malignant), but instead may be required to differentiate among multiple states of truth. For example, a neural network for the differential diagnosis of interstitial diseases was trained to distinguish nine different interstitial diseases,^{6,7} and therefore required nine output units corresponding to the nine diseases.

When multiple outlet units are required, we consider a neural network with four input units, two hidden units and four output units. This structure attempts to distinguish among four different diseases by using four quantitative "symptoms" or "test results" as input data. We used two different sets of four input data distributions, corresponding to the four diseases, to study the relationship between the nature of the input data and the output of the neural network. Each set of the four input data distributions can be represented in four-dimensional input data space in a way similar to the pair of two-dimensional distributions shown in Figs 5A and 5B. The neural network was first trained by four-dimensional input data sets that corresponded to diseases A, B, C, and D with corresponding ideal output values of 1.0 at output units 1, 2, 3, and 4, respectively, and 0.0

at the remaining output units. Thus, output units 1, 2, 3 and 4 represent diseases A, B, C, and D, respectively. Independently sampled input data sets for diseases A, B, C, and D were used for testing.

A generalized form of ROC analysis was used to evaluate the network's performance in differentiating the four diseases. This analysis can be illustrated by the following example. Each input data set arising from disease A, B, C, or D produced values at each of the four output units. To evaluate the performance of the neural network in differentiating disease A from the other three diseases, only the values produced at output unit 1 were considered. At output unit 1, an output value exceeding a fixed threshold level was scored as a true-positive diagnosis if it arose from an input data set produced by disease A, or as a false-positive diagnosis if arose from an input data set produced by any other three diseases. An ROC curve for the detection of disease A was then obtained in the conventional way by calculating TPF and FPF as a function of the threshold level setting. Similarly, ROC curves for the detection of diseases B, C, and D were obtained by considering the values obtained at output units 2, 3, and 4, respectively. Input values for each of the input units were sampled independently from Gaussian distributions with the means for each disease shown in Table 1 and with an SD of 0.2. The experiments were then repeated with Gaussian distributions having an SD of 0.1. Each training and testing sample consisted of 200 and 2,000 data sets, respectively. The average values of A_z obtained in each situation and the estimated SDs of these mean values (σ) also are shown in Table 1.

In the situation shown in the upper portion of Table 1, diseases A and D are unique in the sense that their mean values are the smallest and largest, respectively, at each input unit. Diseases B and C, on the other hand, are somewhat ambiguous in that their mean input values lie between those of diseases A and D at all four input units. As shown in the upper right-hand part of Table 1, the neural network was able to identify diseases A and D well, but it was not able to identify diseases B and C as reliably. To examine the effect of the uniqueness of a disease's input data pattern on its

Table 1. Mean Values of Input Data With Corresponding Truth and Test Results of A_z Values

Disease	Mean Input Values for Training and Testing				Desired Output Values				Test Results (SD = 0.2)		Test Results (SD = 0.1)	
	Input unit 1	Input unit 2	Input unit 3	Input unit 4	Output unit 1	Output unit 2	Output unit 3	Output unit 4	A_z	σ	A_z	σ
A	0.4	0.4	0.4	0.4	1.0	0.0	0.0	0.0	0.866	0.010	0.971	0.002
B	0.5	0.5	0.5	0.5	0.0	1.0	0.0	0.0	0.594	0.043	0.856	0.009
C	0.6	0.6	0.6	0.6	0.0	0.0	1.0	0.0	0.652	0.033	0.803	0.095
D	0.7	0.7	0.7	0.7	0.0	0.0	0.0	1.0	0.888	0.006	0.972	0.002
A	0.4	0.4	0.4	0.4	1.0	0.0	0.0	0.0	0.864	0.014	0.963	0.001
B'	0.3	0.5	0.5	0.5	0.0	1.0	0.0	0.0	0.827	0.027	0.933	0.030
C'	0.8	0.6	0.6	0.6	0.0	0.0	1.0	0.0	0.834	0.014	0.953	0.010
D	0.7	0.7	0.7	0.7	0.0	0.0	0.0	1.0	0.858	0.014	0.963	0.006

Note: Data collected from 200 iterations, 200 training data sets, and 2,000 testing data sets.

detectability, we replaced diseases B and C with diseases B' and C', respectively, as shown in the lower portion of Table 1. In this situation, disease B' and C' provided the smallest and largest mean input value, respectively, at one output unit. As a result, the neural network's performance was greatly improved. From the average A_z values shown in the lower right-hand part of Table 1, it is evident that the network can now differentiate all four diseases. This result suggests that the neural network is able to identify unique characteristics of an input data pattern and can use them efficiently in classifying different patterns. The effect of SNR is also evident in Table 1, where the network's performance is seen to improve for both patterns of input means when the SDs of the input distributions were reduced from 0.2 to 0.1. One might expect that the performance of a neural network depends on both the SNR of each input unit and the extent to which each truth state is associated with a unique pattern of input value means. A comparison of the two situations presented in Table 1 indicates that network performance can be increased either by adding unique characteristics to the input patterns (as in the lower portion of Table 1) or by increasing the SNRs of the existing inputs.

SUMMARY AND CONCLUSIONS

We have studied the effect of both input data and internal structure on the performance of neural networks. Performance of the neural networks was assessed quantitatively in terms of ROC analysis and compared with theoretical calculations. We confirmed that the simple neural networks used in this study can perform

at the level of an ideal statistical decision maker. Furthermore, our simulation study suggests that neural networks possess two potentially useful characteristics: (1) the ability to learn accurately even when the training data are extremely unbalanced with respect to the prevalence of actually positive cases, and (2) the ability to differentiate input data patterns on the basis of their unique characteristics.

Our simulation study was conducted under three idealized conditions: (1) training and testing data provided to each input unit are statistically independent of the data provided to every other input unit; (2) SNRs at all of the individual input units are identical; and (3) the SDs of the Gaussian input value distributions for actually negative and actually positive cases are equal. Those conditions make it possible to focus, as the first step in our investigation, on the fundamental aspects of neural network behavior. We believe that the understanding of the characteristics of neural networks obtained from this simulation study may serve as a useful guide in applying neural networks to a variety of practical data-classification and pattern-recognition problems in medical imaging.

APPENDIX

Figure 8 shows a pair of one-dimensional data distributions. Let (μ_b, σ_b) and (μ_m, σ_m) represent the mean and SD of the distributions associated with benign and malignant cases, respectively. The area under the ROC curve (A_z) is then given by¹⁶:

$$A_z = \Phi \left(\frac{|\mu_m - \mu_b|}{\sqrt{\sigma_m^2 + \sigma_b^2}} \right) \quad (1a)$$

Now consider an N-dimensional Gaussian distribution, $G(\vec{r})$, where $\vec{r} \equiv (r_1, r_2, \dots, r_N)^T$ is a column vector that

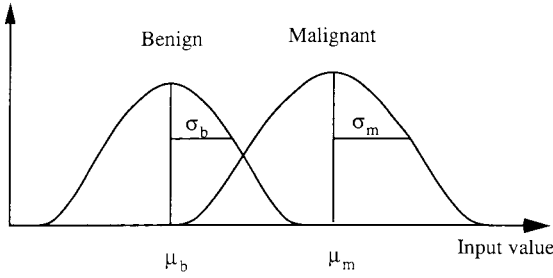


Fig 8. One-dimensional Gaussian distributions with different means and SDs for benign (*b*) and malignant (*m*) cases.

represents the joint outcomes of N random variables with means and SDs represented by the pair of column vectors $(\bar{\mu}_b, \bar{\sigma}_b)$ and $(\bar{\mu}_m, \bar{\sigma}_m)$ for benign and malignant cases, respectively. Any linear combination of the N -dimensional input data that could serve as a decision variable in conventional ROC analysis can be represented generally by the following:

$$z = v_1 r_1 + v_2 r_2 + \dots + v_N r_N = \bar{v} \cdot \bar{r}.$$

The expected values of z for benign and malignant cases are then:

$$E\{z|b\} = \bar{v} \cdot E\{\bar{r}|b\} = \bar{v} \cdot \bar{\mu}_b \equiv \mu_{z|b}$$

$$E\{z|m\} = \bar{v} \cdot E\{\bar{r}|m\} = \bar{v} \cdot \bar{\mu}_m \equiv \mu_{z|m}.$$

The difference of expected values between benign and malignant cases becomes:

$$E\{z|m\} - E\{z|b\} = \bar{v} \cdot (\bar{\mu}_m - \bar{\mu}_b) \equiv \bar{v} \cdot \Delta\bar{\mu} \\ = \mu_{z|m} - \mu_{z|b} \quad (2a)$$

The variance of z for benign cases is given by this equation:

$$\text{Var}\{z|b\} = \text{Var}\{\bar{v} \cdot \bar{r}|b\} = E\{[\bar{v} \cdot (\bar{r} - \bar{\mu}_b)]^2|b\} \\ \equiv E\{[\bar{v} \cdot (\Delta\bar{r})]^2|b\}. \quad (3a)$$

Note that, in matrix notation, the scalar product in the above equation can be written as:

$$\bar{v} \cdot (\Delta\bar{r}) = (\bar{v})^T (\Delta\bar{r}) \\ [(\bar{v} \cdot (\Delta\bar{r}))^2] = (\bar{v})^T (\Delta\bar{r}) (\Delta\bar{r})^T (\bar{v}).$$

We define a covariance matrix C as the expected outer product

$$C = E\{(\Delta\bar{r})(\Delta\bar{r})^T\},$$

whereby the covariance matrices for benign and malignant cases can be written as:

$$C^b = E\{(\Delta\bar{r})(\Delta\bar{r})^T|b\} \text{ and } C^m = E\{(\Delta\bar{r})(\Delta\bar{r})^T|m\}, \quad (4a)$$

respectively. Notice that C^b and C^m are symmetrical matrices. Substituting equation 4a into 3a, we obtain

$$\text{Var}\{z|b\} = E\{[\bar{v} \cdot (\Delta\bar{r})]^2|b\} = (\bar{v})^T C^b \bar{v} \equiv \sigma_{z|b}^2. \quad (5a)$$

Similarly, the variance of z for malignant cases can be

obtained as:

$$\text{Var}\{z|m\} = E\{[\bar{v} \cdot (\Delta\bar{r})]^2|m\} = (\bar{v})^T C^m \bar{v} \equiv \sigma_{z|m}^2. \quad (6a)$$

Notice that $\mu_{z|b}$, $\mu_{z|m}$, $\sigma_{z|b}$ and $\sigma_{z|m}$ here correspond to μ_b , μ_m , σ_b and σ_m , respectively, in 1a. From equations 1a, 2a, 5a, and 6a, we obtain

$$Q \equiv \frac{(\mu_{z|m} - \mu_{z|b})^2}{\sigma_{z|m}^2 + \sigma_{z|b}^2} = \frac{(\bar{v} \cdot \Delta\bar{\mu})^2}{(\bar{v})^T [C^m + C^b] (\bar{v})} = \frac{(\bar{v} \cdot \Delta\bar{\mu})^2}{(\bar{v})^T C (\bar{v})}, \quad (7a)$$

where $C \equiv C^b + C^m$ is a symmetrical matrix.

It is evident from equation 1a that one needs only to choose an appropriate weight vector \bar{v} to maximize Q in order to maximize A_z , because $|\mu_{z|m} - \mu_{z|b}|/\sqrt{\sigma_{z|m}^2 + \sigma_{z|b}^2}$ is non-negative and Φ is a monotonically increasing function. From equation 7a the following can be generated:

$$Q = \frac{(\bar{v} \cdot \Delta\bar{\mu})^2}{(\bar{v})^T C (\bar{v})} = \frac{\left[\sum_{i=1}^N v_i (\Delta\mu_i) \right]^2}{\sum_{i=1}^N \sum_{j=1}^N v_i C_{ij} v_j}. \quad (8a)$$

To maximize Q , we need to find a set of values $\{v_i\}$ such that

$$\frac{\partial Q}{\partial v_i} = 0 \quad \text{for } i = 1, 2, \dots, N.$$

Differentiating equation 8a with respect to v_n ($n = 1, 2, \dots, N$), we obtain

$$\frac{\partial Q}{\partial v_n} = \frac{2(\bar{v} \cdot \Delta\bar{\mu}) \Delta\mu_n \{(\bar{v})^T C (\bar{v})\} - 2(\bar{v} \cdot \Delta\bar{\mu})^2 \sum_{j=1}^N C_{nj} v_j}{\{(\bar{v})^T C (\bar{v})\}^2} = 0 \quad (9a)$$

for $n = 1, 2, \dots, N$, which simplifies to

$$\Delta\mu_n \{(\bar{v})^T C (\bar{v})\} = (\bar{v} \cdot \Delta\bar{\mu}) \sum_{j=1}^N C_{nj} v_j \quad (n = 1, 2, \dots, N).$$

Combining all of these equations for $n = 1, 2, \dots, N$, we obtain, in matrix notation,

$$\{(\bar{v})^T C (\bar{v})\} (\Delta\bar{\mu}) = [(\bar{v} \cdot \Delta\bar{\mu})] C \bar{v}, \quad (10a)$$

which is solved by:

$$\bar{v} = k C^{-1} (\Delta\bar{\mu}) \quad (11a)$$

where k is an arbitrary constant. This weight vector \bar{v} maximizes Q , and therefore A_z . It is interesting to notice, in passing, that the set of weights given by equation 11a is equivalent to that of a "pre-whitening matched filter."^{17,18}

Substituting equation 11a into 8a, we find that the maximum value of Q is given by

$$Q_{\max} = \frac{[(\bar{v} \cdot \Delta\bar{\mu})]^2}{\{(\bar{v})^T C (\bar{v})\}} = (\Delta\bar{\mu})^T C^{-1} (\Delta\bar{\mu}).$$

so from equation 1a:

$$(A_z)_{\max} = \Phi(Q_{\max}) = \Phi(\sqrt{(\Delta\bar{\mu})^T C^{-1} (\Delta\bar{\mu})}). \quad (12a)$$

In the special situation in which the random variables (r_1, r_2, \dots, r_N) are independent, the correlation matrix C becomes diagonal, ie,

$$C = D(\sigma_{mi}^2 + \sigma_{bi}^2), \text{ so } C^{-1} = D(1/[\sigma_{mi}^2 + \sigma_{bi}^2])$$

Substituting this into equation 12a, we find that the maximum A_z for independent input values is given by

$$(A_z)_{\max} = \Phi \left(\sqrt{\sum_1^N \frac{(\mu_{mi} - \mu_{bi})^2}{\sigma_{mi}^2 + \sigma_{bi}^2}} \right).$$

If we define a signal-to-noise ratio for the N-dimensional data distribution as

$$\text{SNR}^{(N)} = \sqrt{\sum_1^N \frac{(\mu_{mi} - \mu_{bi})^2}{(\sigma_{mi}^2 + \sigma_{bi}^2)/2}}, \quad (13a)$$

the maximum attainable value of A_z can be expressed as

$$(A_z)_{\max} = \Phi \left(\frac{1}{\sqrt{2}} \text{SNR}^{(N)} \right). \quad (14a)$$

Finally, we consider a simplified situation in which the SDs on each of the N dimensions are identical, (ie, $\sigma_{bi} = \sigma_b$ and $\sigma_{mi} = \sigma_m$ for $i = 1, 2, \dots, N$) and the absolute values of the separation between the means of the benign and malignant distributions is the same on all N dimensions as

well (ie, $|\mu_{mi} - \mu_{bi}| = \Delta\mu$ for $i = 1, \dots, N$). Then from equation 13a, the N-dimensional SNR becomes

$$\text{SNR}^{(N)} = \frac{\sqrt{N}\Delta\mu}{\sqrt{\frac{\sigma_m^2 + \sigma_b^2}{2}}} = \sqrt{N} \frac{\Delta\mu}{\sigma} = \sqrt{N} \text{SNR}^{(1)},$$

where $\text{SNR}^{(1)} = \Delta\mu/\sigma$ is the SNR for any one input and $\sigma = \sqrt{(\sigma_m^2 + \sigma_b^2)/2}$ is the root-mean-square SD. Substituting into equation 14a, we obtain

$$(A_z)_{\max} = \Phi \left(\frac{1}{\sqrt{2}} \text{SNR}^{(N)} \right) = \Phi \left(\sqrt{\frac{N}{2}} \text{SNR}^{(1)} \right). \quad (15a)$$

According to the arguments presented here, equations 14a and 15a give the maximum values of A_z that can be obtained from a *linear* combination of input values. In general, larger values of A_z may be achievable by nonlinear decision rules. However, one can show that the values of A_z given by equations 14a and 15a are the largest that can be obtained by any decision rule (ie, represent the performance of the ideal decision maker,^{3,18} who uses likelihood ratio as a decision variable) in the special case where $\sigma_{mi} = \sigma_{bi}$ ($i = 1, 2, \dots, N$), which is the case in our simulations.

REFERENCES

1. Hecht-Nielsen R: Neurocomputing. Reading, MA, Addison-Wesley, 1990
2. Pao YH: Adaptive Pattern Recognition and Neural Networks. Reading, MA, Addison-Wesley, 1989
3. Brown DG: Neural networks and ideal observers. Proc SPIE Med Imaging 1090:154-163, 1989
4. Boone JM, Sigilitto VG, Shaber GS: Neural networks in radiology: An introduction and evaluation in a signal detection task. Med Phys 17:234-241, 1990
5. Boone JM, Gross GW, Greco-Hunt V: Neural networks in radiologic diagnosis: I. Introduction and illustration. Invest Radiol 25:1012-1016, 1990
6. Asada N, Doi K, MacMahon H, et al: Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: Pilot study. Radiology 177:857-860, 1990
7. Asada N, Doi K, MacMahon H, et al: Neural network approach for differential diagnosis of interstitial lung diseases. Proc SPIE Med Imaging 1233:45-50, 1990
8. Boone JM: X-ray spectral reconstruction from attenuation data using neural networks. Med Phys 17:647-654, 1990
9. Gross GW, Boone JM, Greco-Hunt V, et al: Neural networks in radiologic diagnosis: II. Interpretation of neonatal chest radiographs. Invest Radiol 25:1017-1023, 1990
10. Clarke LP, Silbiger MS, Naylor WC, et al: Artificial neural net system for interactive tissue classification with MR imaging and image segmentation. 76th Scientific Assembly and Annual Meeting of RSNA, Radiology 177:148, 1990 (abstr)
11. Wu Y, Doi K, Giger ML, et al: Computerized detection of clustered microcalcifications in digital mammograms: Applications of artificial neural networks. Med Phys 19:555-560, 1992
12. Wu Y, Giger ML, Doi K, et al: Artificial neural networks in mammography: Application to decision making in the diagnosis of breast cancer. Radiology, April 1993 (in press)
13. Metz CE: ROC methodology in radiologic imaging. Invest Radiol 21:720-733, 1986
14. Rumelhart DE, McClelland JL (eds): Parallel Distributed Processing. Cambridge, MA, MIT, 1986
15. Metz CE, Shen JH, Herman BA: New methods for estimating a binormal ROC curve from continuously-distributed test results. Presented at the 1990 Joint Meetings of the American Statistical Society and the Biometric Society: Anaheim, CA, August 1990
16. Metz CE: Statistical analysis of ROC data in evaluating diagnostic performance, in Herbert D, Myers R (eds): Multiple Regression Analysis: Applications in the Health Sciences. New York, NY, Am Institute Physics, 1986
17. Wagner RF: Towards a unified view of radiological imaging systems. Med Phys 4:279-296, 1977
18. Wagner RF, Brown DG: Unified SNR analysis of medical imaging systems. Phys Med Biol 30:489-518, 1985