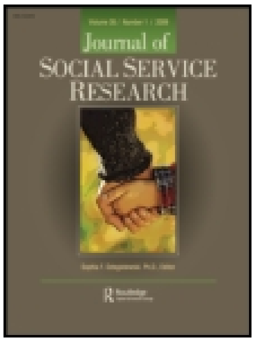


Journal of Social Service Research

ISSN: 0148-8376 (Print) 1540-7314 (Online) Journal homepage: http://www.tandfonline.com/loi/wssr20

Predicting Placement in Foster Care

Thomas P. McDonald PhD , John Poertner DSW & Gardenia Harris PhD

To cite this article: Thomas P. McDonald PhD , John Poertner DSW & Gardenia Harris PhD (2002) Predicting Placement in Foster Care, Journal of Social Service Research, 28:2, 1-20, DOI: 10.1300/J079v28n02_01

To link to this article: https://doi.org/10.1300/J079v28n02_01

Predicting Placement in Foster Care: A Comparison of Logistic Regression and Neural Network Analysis

Thomas P. McDonald
John Poertner
Gardenia Harris

ABSTRACT. In this paper we explore the use of neural network analysis (NNA) as an alternative to logistic regression to predict which children with a founded (indicated) child abuse/neglect report will be subsequently placed in foster care. The main advantages of NNA are that it is a nonparametric technique requiring no assumptions of normality that can readily accommodate both linear and nonlinear relationships and interactions without prior specification by the researcher. The two techniques were found to yield similar classification results for these data; however, NNA provides unique capabilities in analyzing and displaying interactions in predictor variables that may make it more useful for data mining.

KEYWORDS. Foster care, child abuse and neglect, neural network analysis, child welfare

---

Thomas P. McDonald, PhD, is affiliated with The University of Kansas, School of Social Welfare, Lawrence, KS. John Poertner, DSW, is affiliated with the School of Social Work, University of Illinois at Urbana-Champaign. Gardenia Harris, PhD, is affiliated with the School of Social Work, University of Missouri-Columbia.

This paper was supported in part by the Children and Family Research Center, School of Social Work, University of Illinois at Urbana-Champaign which is funded in part by the Illinois Department of Children and Family Services.

INTRODUCTION

Numerous analytical techniques are available to the researcher for developing and testing prediction models. Multiple regression analysis (MRA) is the most common approach when dealing with a continuous outcome variable. Logistic regression (LRA) has become the favored approach when the outcome variable is categorical (Hosmer & Lemeshow, 1989). While not a particularly new technique, neural network analysis (NNA) had largely died out in the 1960s but enjoyed resurgence in the 1980s largely due to the work of John Hopfield. Since that time, the development of new modeling techniques, software and the tremendous leaps in computer processing speed have fueled a growing interest in this approach.

In this paper we compare the results from logistic regression and neural network analyses that attempt to predict which children with a founded (indicated) child abuse/neglect report will be subsequently placed in foster care. We assume that the reader has some knowledge of logistic regression. The emphasis is on the use of neural network analysis, specifically the multi-layer perceptron to address the same analytic problem. We attempt to provide an overview of the technique and the use of Neural Connection software (SPSS) for conducting the analysis.

A BRIEF INTRODUCTION TO NEURAL NETWORK ANALYSIS

Neural network analysis grew from early attempts to use computers to mimic the basic functions of the brain and nervous system. Human and animal nervous systems operate by distributing processing tasks among millions of simple nerve cells called neurons. The neurons receive signals from other cells through input structures called dendrites. These signals are processed and combined by the neurons, and under the right conditions an output signal is sent to the axon. Axons from one cell connect to dendrites of other cells through junctions called synapses. Various electrochemical interactions determine the strength of the signals carried across the organic neural system.

Neural network computing attempts to model this process primarily through software simulations. In these artificial neural networks, inputs (analogous to the dendrites in a biological neural network) pass information to the processing element, also known as a node or neuron. This information is processed by means of some transfer function before being sent on as an output. Each connection has a weight or scaling factor associated with it (analogous to the strength of the synaptic pulse in a biological network). One of the ways that neural network analysis is unique compared to statistical analyses and expert systems is that the values of these weights change in response to each new piece of information that is entered, giving the network the ability to learn (Sarle, 1994).

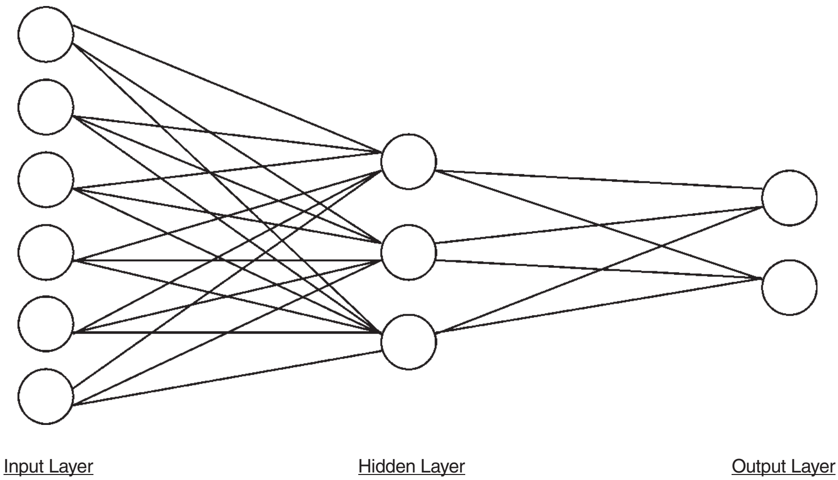
Neural network models can be of two main types: feed-forward and feedback. Feedback networks are much more complex and unstable. Software employing feedback models are not as readily available. The most commonly used neural computing technique is a feed-forward model called the Multi-Layer Perceptron (MLP). The MLP is actually a large family of models that are trained by propagating errors back through the network to adjust the weights used by earlier layers. For this reason, they are also called back propagation or backprop nets.

An artificial neural network model consists of layers of neurons. The basic model consists of three layers: input, middle or hidden, and output. The input layer receives data from the real world, in this case the independent variables of the prediction model. Data is passed from the input layer through the hidden layer(s) where it is weighted, summed and forwarded to the next layer through a transfer function which usually introduces nonlinearity into the neural network model (Garson, 1998, pp. 26-27). This process is shown graphically in Figure 1.

While models can be constructed with more than one hidden layer, most work well with only one and any problem can be represented by using, at most two hidden layers of neurons with nonlinear transfer functions (Bishop, 1995). The number of neurons in the input and output is a function of the variable values. In the simplest case, there is one neuron in the input layer for each predictor variable and one neuron in the output layer for each dependent variable. Input fields may be numeric, text or even pictorial in some software packages (not SPSS). Output fields must be text for decision problems and numeric for prediction problems.

Neural network analysis (NNA) is comparable to several statistical techniques. Its primary application is for prediction and in that context can be seen as an alternative to multiple regression, logistic regression and other forms of the multivariate general linear model. NNA can also be used in classification problems that are commonly analyzed with discriminant analysis and to address problems involving clustering of attributes or cases as one would do with cluster or factor analysis. Comparisons between NNA and these more traditional statistical techniques have found NNA to frequently be superior in terms of predictive power.

FIGURE 1. Simple Multi-Layer Perceptron



While a quite recent phenomenon, NNA use has spread rapidly in the social sciences as documented by Garson (1998). In the area of financial analysis it has become the preferred mode of multivariate investigation. Compared to more standard statistical techniques it has several advantages:

- NNA can accommodate both linear and nonlinear relationships and, unlike standard statistical techniques, it is not necessary to specify the nonlinearities in advance.
- Interactions among the input variables are assumed in NNA and do not have to be specified in advance.
- NNA is a nonparametric technique requiring no assumptions of normality.
- NNA can incorporate a large number and different types of input and output data.
- Neural models can be easily retrained to deal with changes in the environment (i.e., new input data).
- NNA models appear to be robust even when dealing with noisy, overlapping, highly nonlinear and noncontinuous data.

Despite these advantages NNA does face obstacles to its widespread use in the social sciences:

- While neural models lend themselves to prediction, they are much more difficult to evaluate from a causal perspective. Summary statistics and tests of significance that allow the researcher to evaluate the performance of individual predictor variables are not readily available in NNA.
- NNA provides the researcher with a variety of different models and variations to choose from, frequently with no clear-cut basis for decision making. The approach is basically pragmatic. The model that works best is the answer. The search for the best model can be endless.
- Practical limitations hinder its adoption in the social sciences. Software has only recently become more widely available and still lags behind in user friendliness. Much of the literature is outside the social sciences and, while analogous to many statistical techniques, different terminology is used (Garson, 1998).

A CHILD WELFARE APPLICATION OF NNA

Each step in the child welfare continuum of services has major implications for the lives of children and their families and for the costs to the public for these services. Better information about factors associated with each transition step can be useful from several perspectives.

The general philosophy of child welfare services places a higher value on effective early intervention to minimize the degree of intrusiveness and disruption to the family and child. Out-of-home placement is to be avoided if the child can be maintained safely in his/her own home. Numerous studies have attempted to identify predictors of out-of-home placement (Barth, 1994: Catalano, Lind & Rosenblatt, 1999; Children's Bureau, 1997; Fanshel & Shinn, 1978; Groeneveld & Giovannoni, 1977; Lindsey, 1991; Segal & Schwartz, 1985) Better information about which children are at risk of being placed once a report of abuse/neglect has been received can help to identify these children in order to take steps to avoid placement or to better prepare for placement when it is necessary.

The identification of predictors of placement can also point to areas of possible bias that could inform policy and practice changes. For example, African American children are placed at higher rates than Caucasian children, which could point to institutional racism and the need for cultural competency training for workers or better allocation of in-home treatment services. Other identified predictor variables could be more directly causal in their role of either leading to placement (e.g., drug abuse by parents, vulnerability of the child) or promoting family resilience (e.g., family structure, services provided).

Methodology

This study employed a retrospective cross-sectional design using child welfare administrative data. Data were analyzed using Neural Connection 2.1 (SPSS, 1990) and classification results were compared with those obtained from logistic regression also using SPSS (version 9.0). For purposes of the analysis a random sample was drawn from the Illinois Department of Children and Family Services (DCFS), Integrated Database. The focus of the analysis was to predict of out-of-home placement for children with a founded abuse or neglect report.

Database

The Illinois Department of Children and Family Services Integrated Database is constructed and maintained by The Chapin Hall Center for Children at the University of Chicago. This longitudinal relational database was constructed in the late 1980s by acquiring quarterly downloads of data from the Department's administrative data systems. As the database has developed, the quarterly updates continue to be used to update historical data and add new data. The database is constructed from the Department's Child Abuse and Neglect Tracking System and the child placement system called CYCIS. These two systems were developed separately with no common identifier to track children across the systems. For over ten years Chapin Hall staff have developed and refined a probabilistic matching protocol to link child records across the two systems. They use a state of the art matching program and methodology that has been the subject of a great deal of study. Some of these studies have found errors in the matching routine. These errors were subsequently corrected. Probabilistic studies indicated that the match is more than 90% accurate. However, to date it has not been possible to determine the exact accuracy of the match. On less rigorous grounds the match appears to be very accurate since many studies drawing a sample from the database and collecting data from case records have rarely found cases that were mismatched (L. M. Bilaver, personal communication, 1999).

Sample

Children were selected from the Child and Youth Centered Information System (CYCIS). The sample was limited to cases opened in fiscal years 1996 and 1997 that represented the child's first case opening in CYCIS and to children less than 18 years of age at the time of the case opening.

The sample was stratified into two groups. The first group was children with a founded (indicated) child abuse/neglect report who remained at home (intact family). The second group was children with a founded abuse and neglect report who were placed in foster care (placed). Intact family cases were defined as those cases in which the family as a whole received services, and there were no children from that family in state custody at the time of case opening. For children placed out-of-home, only those cases that opened with a first placement in a family foster care or home of relative care were included.

The resulting data set consisted of 4,147 children who were admitted into foster care and 10,135 children who received services in the home. A random sample of 842 cases who were placed in foster care and 964 who received services in their home was drawn for purposes of model testing. This sample was selected to meet the criteria of 95% certainty that the estimated parameter was within 3% of the true value. The criterion for selection of the sample to be used for model replication was increased to 99% certainty that the estimated parameter was within 2% of the true mean, resulting in a sample of 4,946 cases consisting of 2,902 children from intact cases and 2,044 children from child cases.

Predictor Variables

The independent variables for this study were selected by reviewing empirical literature regarding the foster care placement decision. The variables identified in the literature that were available in the database included the following:

- child characteristics (age, gender, race),
- caregiver characteristics (age, gender, race),
- family structure (never married, married, divorced),
- the caregiver and child relationship (biological parent, relative),
- previous allegation history (the number of previous total allegations, the number of previous indicated allegations),
- the severity of the most recent allegation (the number of allegations in the most recent report, the type of allegation identified as most severe in the most recent report),

- case processing variables (initial reporter, the number of investigator home visits, the number of other investigator contacts), and
- DCFS region.

Classification Results

Prior analyses of this data set used logistic regression to identify factors that might influence the decision to place a child (Harris, 1999; Harris & Poertner, 1999). Following a strategy developed by Dattalo (1994), Harris and Poertner (1999) first identified 35 predictor variables that were significant at the .15 alpha level in bivariate analyses with the dependent variable. Forward stepwise logistic regression was then used to select a subset of best predictors and the most parsimonious model. The final logistic regression model derived by Harris and Poertner (1999) is shown in Table 1. Harris and Poertner report that, the twelve variable model shown in Table 1 classify 75.2% of the cases correctly. Eighty-two percent of the intact family cases were correctly classified compared to 68% of the placement cases.

To directly compare the LR and NNA results required that the original data set be split into subsamples. NNA automatically divides the data set into three groups: training data, validation data and test data. The training data set is used to compute model weights. The training data are passed through the model in four stages as the software seeks the best model (defined as the one with the least error). If the training is allowed to proceed long enough the model will eventually “learn” how to correct for all the error in the sample data and to predict perfectly, however, this over-trained model will not generalize to other samples or to the population. To avoid this, the validation data set is used to test the model after each iteration. As the model over-trains, error in the validation data will begin to increase. Training stops when error is optimized in the validation and training sets.

The third NNA data set contains the test data. Test data are set aside and not used in the training of the model. They provide an independent test of the model’s prediction accuracy. While not part of the default settings, SPSS logistic regression software does provide an option to set aside a random test sample. The LR model is then developed on the selected (training) data set and applied to yield classification results for both selected (training) and non-selected (test) cases separately. This option was employed to generate the first two columns under the LR re-

TABLE 1. Logistic Equation

Variable	B	S.E.	Sig	R	Exp(B)
Number of home visits	−1.04	.14	.00	−.15	.35
Number of previous indicated allegations	.28	.05	.00	.10	1.32
Infant	1.33	.15	.00	.17	3.77
Number of contacts	.97	.13	.00	.15	2.65
Number of allegations in the most recent report	−0.89	.15	.00	−.12	.41
Number of previous indicated reports	1.13	.17	.00	.13	3.09
Lack of supervision	.96	.14	.00	.13	2.62
Risk of harm	.71	.14	.00	.10	2.04
Never married	.45	.11	.00	.07	1.56
Toddler (age .5-3 years)	.44	.15	.00	.05	1.56
Nine (9-12 years)	−.58	.20	.00	−.05	.56
Reporter–family and friends	−.45	.16	.01	−.05	.64

sults in Table 2. These can be compared with the classification results for the first two columns under MLP. In both cases a ten percent random sample was used for the test data set (N = 181). The training sample used in the LR model contains both the training and validation samples used in estimation of the NNA model.

In general, any estimation technique will yield better classification results when applied to the data that were used to generate the model weights than it will on subsequent independent samples. This is the case for both the LR and MLP models. The results on the training and test data sets are quite similar for the two procedures. The LR model is better at predicting intact cases than placed cases while the MLP model provides more balanced prediction.

The large administrative data set with which we were working provided a unique opportunity to test the LR and NNA models. The initial sample of 1,806 cases (964 children from intact cases and 842 children from placement cases) was selected to meet the criterion of 95% certainty that the estimated regression parameters were within 3% of the true value. The criterion for selection of the sample to be used for replication was increased to 99% certainty that the estimated parameter was within 2% of the true mean, resulting in a sample of 4,946 cases consisting of 2,902 children from intact cases and 2,044 children from child cases. The third column under each model in Table 2 provides another

TABLE 2. Comparison of Classification Results for Logistic Regression and Multi-Layer Perceptron

	LR			MLP		
	Training (N = 1625)	Test (N = 181)	Replicate (N = 4946)	Training/ Validation (N = 1444 and 181)	Test (N = 181)	Replicate (N = 4946)
Percent Intact Correctly Classified	79%	76%	85%	79%	70%	86%
Percent Placed Correctly Classified	68%	67%	65%	72%	74%	64%
Overall Percent Correctly Classified	74%	72%	76%	76%	71%	77%

independent test of the two models with the much larger replication sample. In both cases we are applying the model weights developed with the training data set to yield predicted outcomes for the new (replicate) sample. The LR and MLP models perform almost identically.

One of the frustrations, or pleasures, of NNA, depending upon the researcher's personal preferences, is that it provides a number of alternatives in constructing a model that may affect the models performance with few clear guidelines for making choices (SPSS, 1990, p.130). These alternatives and the settings used in the above analyses include:

- The numbers of hidden layers–one layer was used in the MLP models.
- The number of nodes in each layer–for both models tested the program generated four nodes in the hidden layer using the automatic node generation option.
- The transfer, or activation, function used by the nodes–the tanh function was used. Sigmoid function can be selected as an alternative nonlinear function and a linear function can also be selected.
- The learning algorithm–conjugate gradient function was used. The steepest descent method is available as an alternative.
- The initial values of the weights between nodes–seed values are initially set at one and can be modified to any value.

All of the MLP models presented here were initially run with default settings. Modifications of these settings did not significantly alter the performance of any of the models.

Interpreting the Models

Social science researchers are generally interested in going beyond classification results from their prediction models to gain some understanding of how the models work. Generally, we wish to know which variables are the most important contributors to the model and whether some variables in the model might be eliminated without reducing the accuracy of prediction. The standard approach to these questions when dealing with logistic analysis is to examine the logistic regression equation table of the type shown above in Table 1. These tables provide several seemingly straightforward ways to evaluate the performance of individual predictor variables in the logistic equation. Regression (b) coefficients cannot be directly compared because they are sensitive to the metric employed in each variable. However, the Wald statistic, based on the ratio of the coefficient to its standard error, can be used to test the significance of each b coefficient. Hauck and Donner (1977) caution against the use of the Wald statistic whenever you have a large coefficient since this situation will lead to an underestimation of the significance of the coefficient. All of the predictor variables in Table 1 are significant at the .01. None of the regression coefficients for these variables appear particularly large thus the significance levels provide one measure of the relative importance of the predictor variables.

The odds ratio (Exp B) is another measure of the "importance" of individual variables in the logistic regression equation. This is a measure of how much the odds of being a member of the represented group (in this case, being placed out-of-home) change with a unit change in the corresponding predictor variable. For example, infants are 3.8 times as likely to be placed as are other age groups of victims of child maltreatment. The statistical significance of an odds ratio can also be evaluated by the Wald statistic. Odds ratios are generally more easily evaluated in a substantive sense, however, there is no guarantee that statistical significance and substantive significance will be in agreement unless sample sizes have been designed with statistical power in mind.

The R statistic is the partial correlation between the dependent variable and each of the independent variables. It can range from -1 to $+1$. The fact that this statistic is labeled a "partial" correlation makes explicit that any measure of the performance of an individual predictor variable in the final equation is always a measure of the unique contribution of the individual variable while controlling for all other variables in the equation. When there is multicollinearity in the predictor variables, this unique contribution of each variable will not be equal to the contribution of a variable by itself nor will the unique contribution of each variable when summed equal the total variance explained by the model. In the absence of carefully selected predictor variables, guided by some theoretical model, it can be extremely misleading to judge the importance of individual predictor variables from the various test statistics provided in the logistic regression equation table. Of course, if multicollinearity does not exist in the predictor variables, these complications do not arise. All three statistics (Wald, R and Exp B) will provide unambiguous tests of the unique and total contribution of each predictor. However, this situation will hardly ever be seen in the real world unless the data were generated from an experimental design. In the real world, the correlation that exists between predictor variables will make it difficult to interpret the contribution of individual predictors regardless of the apparent simplicity of test statistics and associated significance tests.

Neural network models are generally viewed as more difficult to evaluate from a causal perspective or when one is interested in the relative and independent contribution of individual independent variables. This difficulty is inherent in the MLP model as a result of the use of hidden layers, which diffuse the computation of predicted outcomes throughout the model in highly complex ways. This means that there is not a one to one correspondence between weights and predictors. Each predictor is associated with multiple paths to the output and therefore influenced by multiple weights. While formulas have been suggested for combining weights associated with individual predictors (see for example, Schrodt, 1991), no widely accepted procedure has yet emerged and Neural Connection 2.1 provides no such statistics.

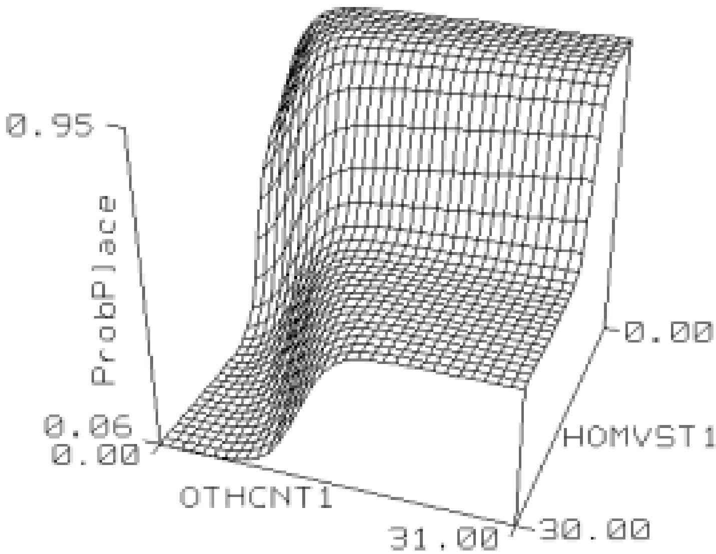
Neural Connection does provide a unique way to understand the relationship between predictors and outcome variables through graphic representations. When the predicted output of a modeling tool is plotted against the values of two input fields, the surface that is plotted is the function that links the input field values and the output of the modeling tool. To overcome irregularities in the distribution of training data values over the range of possible values, Neural Connection, employs the Simulator Tool to create a pseudo test file. In this simulated data file, the values of the input fields are varied over their entire range (or a range of interest to the researcher) while the other input fields are fixed at their mean values (or other values set by the researcher). This plot is another method of model evaluation. Garson (1998) provides this rather loose rule for evaluating the model through its graphic representation: "When the model is working well, this plot will represent a discernible function. When the model is not working well, the plot will not be interpretable."

The need to look at input variables in pairs in NNA is more complicated than statistical or graphic representations that look at simple bivariate relationships with the dependent variables and each independent variable one at a time. However, this also proves to be strength of NNA when the researcher is interested in possible interactions. In addition, the graphic displays generated by Neural Connection do provide creative ways of examining bivariate relationships while holding constant all other input variables. The other major strength of NNA is the flexibility provided for discovering nonlinear relationships. These issues can best be understood through some examples.

Figure 2 is a 3-D wire mesh plot generated by the Graphics Output Tool in Neural Connection that shows the relationship between the number of home visits during the protective service investigation, the number of other contacts made during the investigation and the probability of placement of the child. The most prominent feature of the graph is that this relationship is clearly nonlinear and interactive. Children who have no home visits are highly likely to be placed unless there are also no other contacts, in which case there is a very low probability of placement. If there are some home visits, the probability of placement generally decreases but plateaus around eight visits except in cases where there are few other contacts. More than eight other contacts have little association with placement and below that the relationship is largely negative with more contacts being associated with increased likelihood of placement. Causality here is ambiguous, as it appears that home visits might be reducing or supplanting the need for placement while other contacts may reflect a perceived need for placement on the part of workers.

Figure 3 provides some additional insights into the role of home visits and the possible differential use of home visits for children of different ages. Figure 3 is a plot of the relationship of home visits and child's age as predictors of placement. Again, nonlinearities are evident in this graph. Neural Connection provides a highly useful interactive method for examining these 3-D graphics through the What If? Tool. A screen image of this display for the graph in Figure 3 can be found in Figure 4. On the left side of the What If. . . window is the sensitivity plot of the probability of placement against the age of the child in the vertical axis and the number of home visits in the horizontal axis. This sensitivity plot is representing the same relationship as that shown in the wire mesh plot of Figure 3. However, in the sensitivity plot we are looking straight

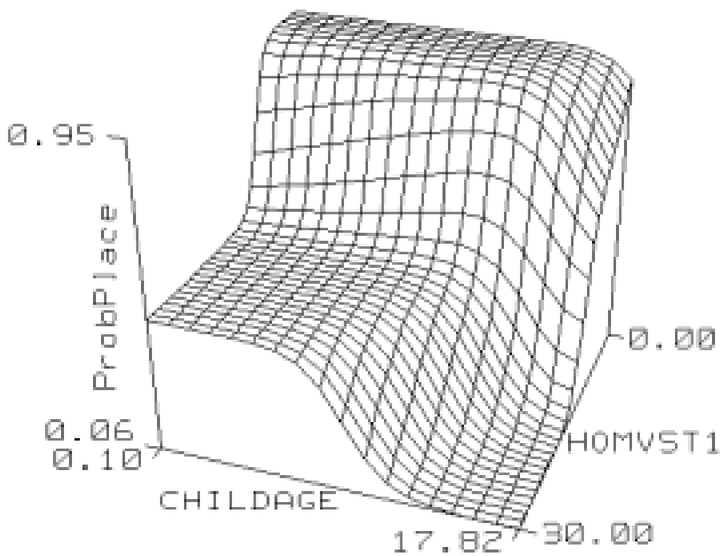
FIGURE 2. Home Visits and Other Contacts as Predictors of Placement



down at the plot rather than from the side as in Figure 3 and the sense of depth reflecting the probability of placement is captured through the use of colors. High output values (in this case high probability of placement) are shown by light colors and low output values by dark colors.

The plot on the right of the What If. . . window is a vertical cross section of the sensitivity plot. The position of the cross section is determined by the position of the moveable horizontal axis in the sensitivity plot on the left. In Figure 4, the horizontal axis has been positioned at roughly one year of age. The plot on the right then provides a graph of the relationship of the number of home visits to the probability of placement for one-year-old children. The user can also position the vertical axis in the sensitivity plot (left side of Figure 4). In Figure 4 we have set the vertical axis at six visits, a point where the cross section plot shows a dramatic break downward. This setting is displayed in the cross section on the right of Figure 4 as the vertical blue line. It provides the base point for the analysis of changes in the home visit variable. The red line in the cross section plot is the change line. By moving it along the cross section you produce a measure of the sensitivity of the output to the inputs that you have chosen. In Figure 4 we have set the red line at the bot-

FIGURE 3. Child's Age and Home Visits as Predictors of Placement



tom of the steep slope where the effects of home visits appear to plateau. The text display at the bottom of the What If. . . window expresses in words the magnitude of the change that is displayed at the fixed values determined by the positioning of the red and blue lines. As we have set these parameters in Figure 4, the text display tells us that the most dramatic change in the probability of placement for one-year-old children occurs between six and thirteen visits where the probability of placement decreases by almost 50 percent from 93 to 48 percent.

By sliding the horizontal axis up we can observe how the relationship between home visits and the probability of placement changes as the age of the child gets older. In performing this operation we found that the relationship remains fairly constant until around age eight years and then begins to transform. Instead of leveling off around thirteen or more visits, the curve gradually begins to resume its steep downward slope as the age of the children increases beyond eight. In Figure 5 we have reproduced the What If. . . window with the child's age at approximately 12 years old. The cross section plot on the right of Figure 5 shows the same dramatic drop in the probability of placement between six and

FIGURE 4. Sensitivity Plot–Child's Age and Home Visits as Predictors of Placement (Child's Age Set to 1)

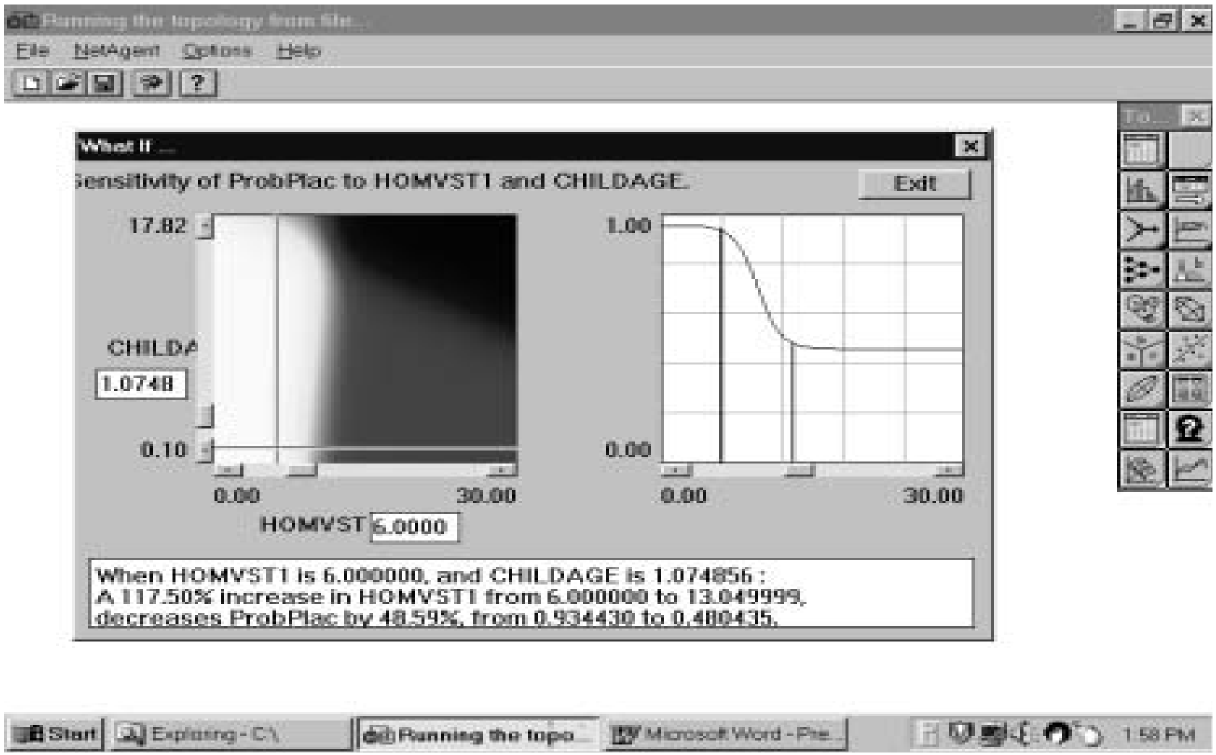
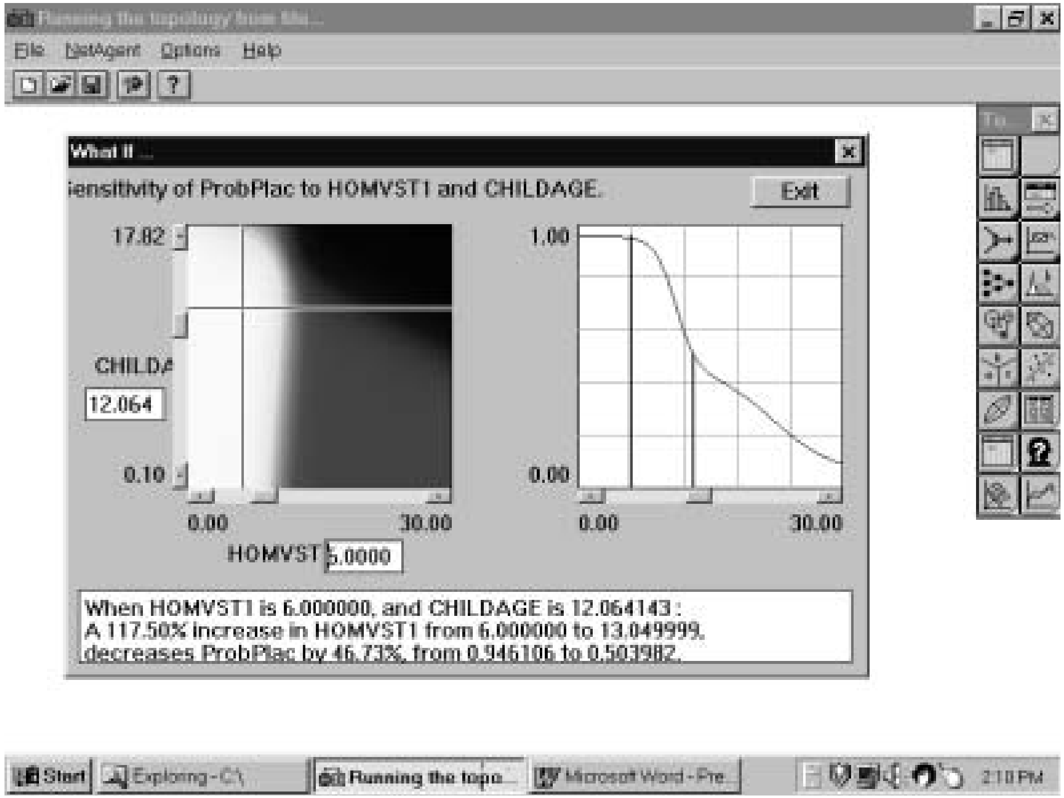


FIGURE 5. Sensitivity Plot–Child's Age and Home Visits as Predictors of Placement (Child's Age Set to 12)



thirteen visits. However, instead of leveling off after 13 visits the decline continues at only a slightly less steep descent.

CONCLUSIONS AND DISCUSSION

Classification results from the logistic regression and neural network analysis were found to be comparable for the model tested here and for the specific data set employed. Both models correctly classified between 70 and 80 percent of the cases. Despite the failure of NNA to yield better classification results in this application, there remain several reasons for preferring it to LR. The fact that NNA is a nonparametric method, requiring no assumptions of normality is a comfort to those researchers who know that this assumption is seldom justified. When operating without the safety of a well-reasoned theory, it is also a comfort and a highly practical convenience to be assured that interactions or nonlinearities will be found without prior specification. Nevertheless, the results from this particular application would also support the position that LR is fairly robust and, for certain models, more complex nonlinearities and interactions may not be that important if one's goal is simply prediction. In addition, LR appears to provide superior tools for achieving parsimony in a model. Forward and backward selection techniques, which are widely used in regression based techniques, are not currently available in NNA.

On the surface it might appear that LR provides more readily interpretable results regarding the contribution of individual variables and an explanation of how the model works. We would maintain that this is largely an illusion unless the world of the researcher, is composed of largely independent events. In multiple regression analysis with correlated independent variables, all the common statistics measuring the strength of independent contributions of individual independent variables are actually measures of highly complex partial correlations. If one is fully aware of and fully acknowledges the conditional nature of these relationships they no longer appear as easily interpretable.

The graphic representations from the NNA presented here more accurately capture the complexity of these relationships when they exist. While considerably more complex than a Wald statistic with an associated significance level, they do provide a type of summative power that becomes more obvious as the researcher gains experience working with the graphs. While it may be true that "a picture is worth a thousand words," it will still be necessary for the researcher to provide a narrative story about what can be found in the picture. This story will be more complex and lengthy than a single statistic but also more accurate.

The situation changes when the researcher is able to operate from a theoretical perspective that allows testing of a specific model. In this case regression based techniques (path analysis and structural equation modeling with and without latent variables) would appear to be preferable.

The Adoption and Safe Families Act places an emphasis on developing and reporting child welfare outcomes that require states to collect, maintain and report data to the Department of Health and Human Services. These requirements in addition to federal support for the development of child welfare management information systems (e.g., SACWIS) are producing a large number of state specific child welfare databases. Data analysis or data mining with these databases provides an opportunity to obtain significant insights into child welfare decision making. However, the data in these systems is generally defined and collected for administrative purposes rather than research. This presents significant data analysis challenges. This study compares logistic regression with its accompanying assumptions about the data that may not be met by administrative data to neural networking analysis that is often considered more appropriate for data mining. The results of this study suggest that there may not be as many differences between these data analysis procedures and that each has its own advantages and disadvantages for use in this environment.

The purpose of this paper was to explore the utility of two alternative procedures for prediction in the social sciences. Any conclusions drawn must be tempered by considerations as to their potential lack of generalizability beyond the data employed here. A second intended function of this paper was to introduce the reader and the field of social work researchers to a new data analysis procedure that has not had wide usage in our field. It is hoped that this article will facilitate and encourage additional attempts to apply these techniques and to report comparisons with more alternative approaches. It is only through this increased experience that the field will be able to draw conclusions about the utility of neural network analysis and to grow more sophisticated in its application.

REFERENCES

Barth, R. P., Courtney, M., Berrick, J. D., and Albert, V. (1994). *From child abuse to permanency planning: Child welfare services pathways and placements.* New York: Aldine De Gruyter.

Bishop, C.M. (1995). *Neural Networks for Pattern Recognition.* Oxford: Oxford University Press.

Catalano, R. A., Lind, S. L., and Rosenblatt, A. B. (1999). Unemployment and foster home placements: Estimating the net effect of provocation and inhibition. *American Journal of Public Health*, *89 (6)*, 851-855.

Dattalo, P. (1994). A comparison of discriminant analysis and logistic regression. *Journal of Social Service Research*, *19*, 121-144.

Department of Health and Human Services, Children's Bureau. (1997). *National study of protective, preventive, and reunification services delivered to children and their families.* Washington, DC: U.S. Government Printing Office.

Fanshel, D., and Shinn, E. B. (1978). *Children in foster care: A longitudinal investigation.* New York: Columbia University Press.

Garson, G.D. (1998). *Neural Networks: An Introductory Guide for Social Scientists.* Thousand Oaks, CA: Sage Publications Ltd.

Groeneveld, L. P., and Giovannoni, J. M. (1977). Disposition of child abuse and neglect cases. *Social Work Research and Abstracts*, *13*(2), 24-30.

Harris, G. (1999). *What is the place of race in foster care placement decisions?* Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign.

Harris, G., and Poertner, J. (1999). *Factors that influence the decision to place a child.* Urbana, IL: Children and Family Research, Center School of Social Work, University of Illinois at Urbana-Champaign.

Hauck, W.W., and Donner, A. (1977). Wald's Test as applied to hypotheses in logit analysis. *Journal of the American Statistical Association*, *72*, 851-853.

Hosmer, D.W., and Lemeshow, S. (1989). *Applied Logistic Regression.* New York, NY: John Wiley and Sons.

Lindsey, D. (1991). Factors affecting the foster care placement decision: An analysis of national survey data. *American Journal of Orthopsychiatry*, *61*, 272-281.

Sarle, W.S. (1994). Neural networks and statistical models. In *Proceedings of the Nineteenth Annual AS Users Group International Conference.* Cary, NC: SAS Institute.

Schrodt, P. A. (1991). Prediction of interstate conflict outcomes using a neural network. *Social Science Computer Review*, 9(3), 359-380.

Segal, U. A., and Schwartz, S. (1985). Factors affecting placement decisions of children following short-term emergency care. *Child Abuse and Neglect*, *9*, 543-548.

SPSS Inc./Recognition Systems Inc. (1997). *Neural Connection 2.0 User's Guide.* Chicago, IL: SPSS Inc./Recognition Systems Inc.

Received: 07/00
Revision Received: 02/01
Accepted: 03/01