From the editor .................................................................................................................................................. 2

Adaptive Fuzzy Regression Model for the Prediction of Dichotomous Response Variables using Cancer data: A Case Study ......................................................................................................................... 3

The development of a web-based platform for fathers of children with Autism .................. 11

Does the availability of audio podcasts enhance the classroom experience for first year dental students? Data on use and perceived benefits ................................................................. 16
Welcome to the April 2008 issue of Learning Technology.

This newsletter focuses on bringing emerging technologies in education to the readers. New developments and practices with learning technologies are the core of this newsletter. This issue covers adaptive technique for predicting dichotomous responses to educating parents of autistic children through Internet.

Nagar and Srivastava proposed an adaptive technique in the prediction of dichotomous response variable by combining fuzzy concept with statistical logistic regression which was tested on cancer dataset in predicting cancer susceptibility. Ferdig et. al. looks at the introduction of an Internet based system to educate fathers of autistic children in acquiring the requisite skills. They then identify the impact of this technology-based intervention on fathers and the results effects on their children with autism. Whitney and Pessina recorded their lectures and provided them to their students and investigate the effects of these podcasts on their learning. They looked at the possibility of increased active learning with the availability of podcasts and how podcasts supplement the learning content.

This newsletter focuses publishing new and emerging technologies in education focussing on advanced learning technologies and its usage in different contexts. Please feel free to bring forward your ideas and views.

Besides, if you are involved in research and/or implementation of any aspect of advanced learning technologies, I invite you to contribute your own work in progress, project reports, case studies, and events announcements in this newsletter. For more details, please refer author guidelines at http://lttf.ieee.org/learn_tech/authors.html.

Ali Fawaz Shareef, PhD
Director General
Centre for Open Learning
Maldives
a.f.shareef@ieee.org
Adaptive Fuzzy Regression Model for the Prediction of Dichotomous Response Variables using Cancer data: A Case Study

Abstract
This paper proposes an adaptive technique in the prediction of dichotomous response variable by combining fuzzy concept with statistical logistic regression. The model was tested on cancer dataset in predicting cancer susceptibility. In this paper we will present the development, evaluation and validation of the proposed model based on the experiment carried out. Explanatory power of the adaptive model was calculated and compared with fuzzy neural network and statistical logistic regression models using calibration and discrimination techniques. Area under ROC values calculated indicates that the proposed model has compatible predictive ability to both fuzzy neural network and statistical logistic regression models.

Keywords: Artificial Neural Network, Fuzzy Regression, Logistic regression

Introduction

Precise and accurate predictive models are very important in screening initiatives. The need for new approaches and philosophies in modeling cancer prediction and susceptibility are influenced by the recent advances in soft computing as well the questionable accuracy and inapplicability to individual prediction of previously sought after statistical analysis techniques. Thus establishing precise predictive models become increasingly more difficult for multivariable predictive models. Traditionally, such regression problems have been addressed by statistical logistic regression techniques for binary dependent variables.

Machine learning technique and interval

Prediction

A machine learning technique is an algorithm that estimates an unknown dependency between a set of given input variables and its response variable. When such dependency is discovered, it can be used to predict or deduce the future output associated with a different set of input values. This is done by identifying the target function that best describes the behavior governing the input-output pattern. Learning in this context refers to the process of minimizing the difference between observed data and model output [7].

An interval prediction is usually comprised of the upper and lower limits between which a future unknown value is expected to lie with a prescribed probability. The prediction interval deals with the accuracy of the estimates with respect to the observed target values [7]. The use of prediction interval in machine learning is appropriate when dealing with multivariate functions where available data are very imprecise and limited and when explanatory variables are interacting in uncertain, vague manners [1]. In other words a fuzzy phenomenon is best modeled by a fuzzy functional relationship. The use of prediction interval in machine learning is referred to as fuzzy linear regression technique.
Motivation of study

Existing Prediction Techniques include statistical techniques and artificial intelligent techniques like Artificial Neural Network (ANN), Support Vector Machine (SVM), Fuzzy Logic, k-nearest neighbors (k-NN), Fuzzy Neural Network (FNN) and Fuzzy Regression [9, 10]. However there are limitations and drawbacks of the above listed prediction techniques.

Problems normally arise in statistical prediction when there is an inadequate number of observations and when distribution assumptions are not satisfied [1]. As for the artificial intelligent prediction techniques, common limitations involve low interpretation ability due to the “Black box” nature of the model (ANN and SVM), limited model ability to explicitly identify possible causal relationships between variables, over fitting problems (ANN and SVM), difficult to build (k-NN), lack in flexibility to incorporate new knowledge (SVM), risk of eroding old but valid information when new knowledge are introduced in the system (SVM) and unsuitable use for high-dimensional data (SVM) [6, 9, 10].

Thus the main question that sparked this study was whether there exist new measures particularly among the artificial intelligent techniques that can be used in predicting binary outcome. The proposed model was supposed to provide answers to the following research questions:

• How do we improve the prediction accuracy using artificial intelligent techniques?
• What can be used to handle ambiguous relationship between the independent (explanatory) and dependent (response) variables?
• What can be introduced in the prediction of dichotomous outcome?
• How do we analyze the non-linear relationship between the independent and dependent variables in multivariate environment?

As a result, an adaptive model was developed by combining the concept of fuzzy with statistical logistic regression. New algorithm to be used for intrinsically linear functions involving linear transformation processes was formulated. This adapted fuzzy logistic regression model can then be used to deduce prediction interval output for binary response variable.

This paper is organized as follows: section I gives the introduction of the proposed fuzzy regression model, section II describes the theory that underlies fuzzy linear regression and fuzzy logistic regression predictive models. The algorithm adapted is shown in section III. Section IV discusses the experiment conducted and model validation. Finally in section V, conclusions from the presented work are drawn.

Underlying theories for the adaptive fuzzy logistic regression model

Fuzzy linear regression theory

Regression analysis is an estimation method used in finding a crisp relationship between the dependent and independent variables and also used to estimate the variance of measurement error. Fuzzy regression analysis is an extension of the classical regression analysis in which some elements of the models are represented by fuzzy numbers [3]. Fuzzy regression methods
have been successfully applied to modeling problems in financial forecasting and engineering [2, 8, 11].

There are two categories of fuzzy regression analysis; the first is a possibilistic regression analysis which is based on possibility concepts. Possibilistic regression analysis uses fuzzy linear system as a regression model whereby the total vagueness of the estimated values for the dependent variables is minimized. It was first proposed by Tanaka et al. [1, 3].

The second category of fuzzy regression analysis adopts the fuzzy least squares method (FLSM) for minimizing errors between the given outputs and the estimated outputs. The advantage of Tanaka’s possibilistic model is in its simplicity in programming and computation, while FLSM in its minimum degree of fuzziness between the observed and estimated values [3].

**Statistical logistic regression theory**

Logistic regression is a mathematical modeling approach that is used to describe the relationship between several explanatory variables X’s to a dichotomous dependent variable Y [5]. Logistic regression can be used to predict the outcome from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these. That is, logistic regression makes no assumption about the distribution of the independent variables. They do not have to be normally distributed, linearly related or of equal variance within each group. The dichotomous dependent variable can take the value of 1 with a probability of success P, or the value of 0 with probability of failure 1-P. This type of variable is called Bernoulli (or binary) variable. The relationship between the predictor and response variables is not a linear function in logistic regression, instead, logistic regression function is used which is the logit transformation of P[5]

\[
\ln\left(\frac{P}{1+P}\right) = \sigma + b_1 x_1 + b_2 x_2 + \ldots + b_j x_j
\]

\[
P = \frac{1}{1 + e^{-(\sigma + b_1 x_1 + b_2 x_2 + \ldots + b_j x_j)}}
\]

where P is the probability of a 1, e is the base of the natural logarithm (about 2.718) and a and b are the parameters of the model.

**The adaptive fuzzy logistic regression model**

The adaptive fuzzy logistic regression model is based on Tanaka’s possibilistic regression analysis described above in which the response variable Y is written as

\[
Y = A_0 x_0 + A_1 x_1 + A_2 x_2 + \ldots + A_j x_j + \ldots + A_k x_k
\]

where Y is the fuzzy output, \( x = [x_1, x_2, \ldots, x_k]^T \) is the real-valued input vector of independent variables and each regression coefficient \( A_j, j=0, \ldots, k \), was assumed to be a symmetric triangular fuzzy number with center \( \alpha_j \) and half-width \( c_j, C_j \geq 0 \) [3,4].
Tanaka’s possibilistic fuzzy regression technique is however applicable to linear functions only [4]. Due to the fact that binary response variable defies the linearity functional relationship that must be satisfied, suitable transformation involving logit (logarithm of odds) transformation must be carried out to unfold the hidden linear relationship. Data must be pre-processed before being fed into a possibilistic fuzzy linear regression model which is then solved by using linear programming to produce a set of corresponding output in an interval form. The output represents the logarithm of the odds for the event to occur. Finally the output is transformed back into the probability of the event occurring by inverting the logarithm of the odds (logit) values. In the algorithm presented here it is assumed that the logarithm of the odds (logit) is linearly related to X’s, the independent variables after undergoing the logit transformation. The algorithm for the adaptive model is summarized in the diagram below:

Experimental application, model validation and results interpretation

The adaptive fuzzy logistic regression model has been tested on a sample set consisting of 150 cancer patients and 136 controls data provided by the Mahatma Gandhi Hospital and research center, Jaipur. This is to illustrate the feasibility of the adaptive fuzzy logistic regression algorithm in predicting oral cancer susceptibility. The data set $x = [x_1, x_2, \ldots, x_n]^T$ refers to the input variables consisting of demographic factors (age group, gender, ethnicity group), risk habits associated with oral cancer (cigarette smoking, alcohol drinking, tobacco chewing) and molecular markers (Gstm1 and Gstt1). The choice of input set is determined based on literature search and discussion with cancer experts from the Mahatma Gandhi Institute. The response variable is binary in nature describing the health status of either having oral cancer or healthy. A total of 17 different sets of input variables were experimented including:

- Set 1: Tobacco-Chewing habit
- Set 2: Cigarette Smoking habit
- Set 3: Alcohol-Drinking habit
- Set 4: Risk Habits (Smoking, Chewing & Drinking)
- Set 5: Gstm1 molecular marker
Set 6: Gstt1 molecular marker
Set 7: Molecular markers (Gstm1 & Gstt1)
Set 8: Molecular markers & Risk habits
Set 9: All markers (risk habits, molecular, age group, ethnic group & gender)
Set 10: Age group
Set 11: Ethnic group
Set 12: Gender
Set 13: Chewing, Age, Ethnicity & Drinking
Set 14: Chewing, Age, Ethnicity & Smoking
Set 15: Chewing, Age, Ethnicity & Gstm1
Set 16: Chewing, Age, Ethnicity & Gstt1
Set 17: Chewing, Age, Ethnicity & Gender

For validation purposes, fuzzy neural network and classical statistical logistic regression models were constructed and the same 17 input data sets were fed into them to predict oral cancer susceptibility. The fuzzy neural network (FNN) model was constructed by combining the learning capability of neural networks with the expressiveness of fuzzy if-then rules using linguistic variables to produce fuzzy neural network models ANFIS. The ANFIS system pioneered by Jang (1992) posses the main component of fuzzy inference system including fuzzification, implication and defuzzification.

The predictive performances for all the three models were assessed using Receiver Operating Characteristic (ROC) curves. The results are presented below:

![Figure 1: ROC curves for Fuzzy Neural Network Model](image1)

![Figure 2: ROC curves for the Adaptive Fuzzy Regression Model](image2)
Comparison of the tabulated AUC values for the three ROC curves shown above suggests that the adaptive fuzzy regression model has compatible ability to fuzzy neural network model and classical statistical model in predicting oral cancer susceptibility based on significantly similar AUC values for all the models.

Similarly, the adaptive fuzzy logistic regression models exhibits similar variable selection ability as the other two models. This is reflected in the ranks of the AUC values for the different contributing factors. For this particular data set, the AUC values for drinking, chewing & smoking decreases (in that order) suggesting that among the three risk habit factors, drinking has the best predictive power followed by chewing and smoking in predicting oral cancer susceptibility (AUC values of 0.707, 0.661, 0.441) for all models.
AUC values of ROC curves for the three models indicate that GSTM1 and combination of GSTM1 with GSTT1 have the least predictive ability since they all show the lowest AUC values of zero. The three models also suggest that the variable set consisting of chewing habit, ethnic group, age group and drinking habit exhibits the highest predictive ability (AUC values of 0.904, 0.897 & 0.852) hence can be considered as the ‘optimal’ variable set for the prediction of oral cancer susceptibility.

**Conclusion**

The possibilistic fuzzy linear regression introduced by Tanaka is being adapted in this study to produce the adaptive fuzzy logistic regression model. This adaptive model was experimented on an oral cancer data set to determine the association between a set of explanatory variables and its corresponding dichotomous response variable. The algorithm formulated can be generalized into prediction problem involving other types of intrinsic linear functions in the fuzzy environment with suitable linear transformation process. The comparatively good results obtained in this application suggest that the adaptive fuzzy logistic regression approach is reasonable, desirable and effective in producing a valid and transparent intelligent exploratory predictive model with dichotomous response variable.

**References**


Pankaj Nagar
Asst. Professor, Department of Statistics
University of Rajasthan
Jaipur – 302015 (Rajasthan)
India
pnagar121@gmail.com

Sumit Srivastava
Research Scholar, Department of Statistics
University of Rajasthan
Jaipur – 302017 (Rajasthan)
India
sumit.310879@hotmail.com
The development of a web-based platform for fathers of children with Autism

Introduction

Autism, once a rare developmental disability, appears to have dramatically increased in recent years. Once occurring in five out of every 10,000 births, recent research indicates that 1 of every 150 children is diagnosed with autism (http://www.autism.com/). While there are a wide variety of treatment options for autism including educational and behavioral interventions, medications, and therapies, some may lead to great improvement while others may have little or no effect (http://www.autism.org).

One area of research that has shown increasing relevance to treatment is the role of family involvement in child development. For instance, researchers examining the role of parents suggest that home intervention programs can help autistic children function more independently in the community as adults (Ozonoff & Cathcart, 1998). Parent participation can also facilitate generalization of the child’s learned skills over a variety of settings (Levy, Kim, & Olive, 2006). There are also reports of children in intervention programs with parental involvement who had increased intelligence scores, which in turn improved their ability to participate in general education (Levy, Kim, & Olive, 2006).

Much of the research on parental involvement has come from looking at mothers as the primary caregivers. Recent work, however, has focused on fathers and the effect of their involvement on child development (Lamb, 1987; Tiedge and Darling-Fisher, 1996). In our own in-home work with fathers, we have provided evidence that that fathers acquired and successfully implemented the training skills they were taught and that the autistic children exhibited improvement in their precommunication skills (Elder, Valcante, Won & Zylis, 2003; Seung, Ashwell, Elder & Valcante, 2006). Additionally, when fathers of children with autism are involved in their child’s life, they report increased feelings of parental competence and marital satisfaction than fathers who are not as involved (Seung, Ashwell, Elder & Valcante, 2006). In sum, results of our research indicate that the in-home training for fathers of children with autism is effective and valued by the participating families.

One of the obvious challenges of our work is that it requires in-home visits. This becomes potentially problematic at two levels. First, we have a successfully-demonstrated intervention; however, if the father does not acquire all of the requisite skills, or the father’s skill level remediates over time, it requires another at-home visit for re-training. Second, success should equate to the delivery of this intervention to multiple households; staffing issues prevents such issues. Limited research has been conducted to study the relationship between technology and father involvement; therefore, a novel system has been built to study the impact of a technology-based intervention on fathers and the resulting effects on their children with autism.
Description of the system

The Autism Family Training Project is a web and database driven technology that allows researchers and parents to interact online. The main website (http://autism.coe.ufl.edu; Figure 1) permits anyone from the public to learn more about the project and related resources. It also allows researchers and parents to log-in.

Figure 1: Screenshot of the homepage.
Once parents have logged in, they have the opportunity to participate in a number of ways (Figure 2). First, they have access to recent news and to a calendar of upcoming events (e.g., training). Second, they have access to instructional videos. The instructional videos include actual training as well as exemplary practices related to the instruction. Third, parents can see examples of personal videos. When researchers visit homes, they take video of the fathers interacting to show them personal examples of behaviors that need to be reinforced or changed. (A long-term goal would be to have the fathers actually upload their own videos or to login using a webcam.) When a father views a video, he can take notes about the video to remember what went well or needed to be changed that day. He can also take notes so that later, when he shows the video to the mother (or siblings, grandparents, and other family members), he can remember various aspects of the training. A resource library and discussion forum (currently between the father and researchers) provides additional support.
The researcher can also log-in to the site. In addition to uploading videos and checking on parent access, researchers can interact with families through discussion forums and calendars. These features will be particularly important in the use of the system for those without in-home visits.

**Conclusion**

Currently, the system is being piloted only with fathers who have received in-home training. One goal of this work is to determine whether fathers who receive in-home training are able to be re-trained using only a web-based intervention. A second goal is to determine whether fathers use this system to scaffold their interaction with other family members (e.g. mothers, siblings). With a deeper understanding of these issues, a final goal will be to determine whether this type of training could benefit new fathers using only a web-based intervention.
References


Ferdig, R.E.
University of Florida
rferdig@ufl.edu

Amberg, H.G.
University of Florida

Elder, J.
University of Florida

Valcante, G.
University of Florida

Donaldson, S.
University of Florida

Bendixen, R.
University of Florida
Does the availability of audio podcasts enhance the classroom experience for first year dental students? Data on use and perceived benefits

Abstract

Lectures in Anatomical Sciences-I, a didactic course covering topics in histology and neuroanatomy, were audio recorded and made available to students in the fall of 2006 at Boston University’s Goldman School of Dental Medicine. To assess audio recording usage and contributions to the learning process, a questionnaire was developed in collaboration the Office of Educational Research, also at the Goldman School of Dental Medicine. The questionnaire was administered to students with the standard course evaluation at the completion of the course. There was a 78% response rate to the survey (90/115). Of the students responding, 56% reported using the lecture recordings. Data reveal that 56.8% of the students who used the recordings listened to the lectures within one week and that the majority of students listened to lectures in their entirety. When asked to the respond to the statement “my learning was enhanced by the use of the lecture recordings,” 93% of users chose either “agree” or “strongly agree.” Additionally, students indicated that the lecture recordings offered the opportunity to actively engage and participate during class. This technology-based resource may increase active learning for all student users and provide an important supplement for dedicated students committed to learning course content.

Introduction

The curriculum for first year dental students is intensive, with over 20 hours of weekly lecture-based instruction and additional extended laboratory sessions. In class, students typically focus their attention on lecture content while taking notes and synthesizing the presented material. The volume of material can be overwhelming and students may leave lecture with concerns regarding the accuracy of their note taking and understanding of key concepts. This was the case during the fall of 2006 when class officers representing the first year dental students (DMD-1), at Boston University’s Goldman School of Dental Medicine, requested that all lectures in Anatomical Sciences-I be audio recorded. Anatomical Sciences-I is a didactic course that covers the topics of histology and neuroanatomy. At the time of the request, the students felt strongly that the opportunity to review an audio recording of lectures would enhance their proficiency with course content. The Office of Information Technology (IT) at Boston University’s Goldman School of Dental Medicine supported the students’ request; members of this Department provided the technical resources for this pilot project. Prior to the formal implementation of this pilot program, several students, with the permission of faculty, were recording lectures for use as a personal accessory learning tool. The introduction of the audio recordings on the password protected, school managed website allowed students without this technology to access lecture recordings. Given the potential educational benefits of this program, the Course Directors felt that a formal assessment of audio recording usage and contributions to the learning process was indicated.
Materials and Methods

At the Goldman School of Dental Medicine, laptop computers are distributed to all DMD-1 students during orientation and students are oriented to the Blackboard CourseInfo™ website, which is used to distribute teaching materials. Our students can easily navigate the password protected Blackboard CourseInfo™ website and are generally comfortable downloading course materials. PowerPoint® files of Anatomical Sciences lectures are routinely made available in CourseInfo™ folders. At student request, audio recordings of corresponding lectures were added to the Blackboard CourseInfo™ website in separate folders, with only registered students granted access. The addition of audio recordings posed no technical issues or concerns. The audio recordings were obtained using Audacity® software and were posted as downloadable MP3 files.

To assess the use of lecture recordings, a questionnaire was developed in collaboration the Office of Educational Research at Boston University’s Goldman School of Dental Medicine. The questionnaire was administered to DMD-1 students with the standard Anatomical Sciences-I course evaluation at the completion of the course. The questionnaire included both open-ended questions that allowed students to describe their experiences with the audio recordings as well as closed-ended questions in which students selected from a dichotomous scale (yes/no) or ordinal scale (1-5).

Results

There was a 78% response rate to the survey (90/115). Of the students responding, 56% reported using the lecture recordings. Although the audio recordings can be easily downloaded onto a portable device, a computer was the primary tool used to listen to the lecture recordings (laptop: 84%; PC 12%). Few students (4%) reported using a mobile device, such as an iPOD.

The data also revealed that 58% of students using the recordings listened to the lecture within one week, with 14% listening 1-2 days after the lecture, 26% listening 3-4 days after the lecture and 18% listening 5-7 days after the lecture (Fig. 1). Thirty-four percent of students listened just prior to the examination (Fig. 1).
Most students listened to lectures in their entirety; 30% listened to all lectures in their entirety and 46% listened to selected lectures in their entirety (Fig. 2).

When asked if the availability of the audio recordings changed in-class note taking strategies, many students indicated that this resource provided the opportunity to focus on “listening to” and “understanding” course content rather than “worrying about writing down details.” In addition, most students responded that they “strongly agree” that their learning in the course was enhanced by the use of lecture recordings (Fig. 3). Specifically, when asked to respond to the statement “my learning was enhanced by the use of the lecture recordings,” 93% of users chose either “agree” or “strongly agree.” Several students expanded on this response, indicating that the lecture recordings offered an opportunity better understand the “big picture.”
Finally, of the 42% of the students who reported that they did not use the recordings, the most common reason reported was “lack of time.”
Discussion

Technology is moving into the educational setting. Courses with a web-based component are part of the educational experience for many students. In fact, the technology of podcasting lectures was recently pioneered at the University of Michigan Dental School in 2004. The term “podcasting” stems from the words iPod and broadcasting. Podcast technology allows the posting of downloadable audio files. This is in contrast to select individuals audiocassette recording for personal use or University sponsored audiocassette recording that may require students to listen to lecture recording in a library setting or may limit the time that an audiocassette may be removed from the library. Since the initial introduction of podcasting, its usage and effectiveness has not been extensively explored in the literature, but poses an interesting array of questions regarding student learning or perception of learning.

At the Goldman School of Dental Medicine, both local and international students are represented in the DMD-I class; these students come with diverse cultural and academic backgrounds that may shape students’ learning needs. In addition, a variety of learning styles have been described and studied, with research demonstrating that the medium through which course content is disseminated may impact learning. The addition of new technologies, such as easy access to lecture audio recordings, provides students with an additional educational resource that supplements traditional didactic lectures.

In addition to providing a mechanism for review of lecture content outside the classroom, the availability of lecture recordings may also enhance the in-class experience. Of the students who listen to the audio recordings, 93% believe that their learning was enhanced by the use of this resource. Although the anonymous nature of the process did not allow us to compare exam scores between users and non-users, the users indicated that, irrespective of examination scores, they felt that they gain a greater appreciation and understanding of the “big picture.” As educators of adult learners, our role is to facilitate an active learning environment. By providing a resource for students to review lectures for specific details, their attention in-class can be diverted to lecture participation. Cognitive researchers suggest that engagement in the active learning process enhances learning and improves recall.

In summary, this technology-based resource may increase active learning for all student users and provide an important supplement for the struggling, yet dedicated student. In fact, the greatest benefit may be to students who have difficulty understanding course content within the confines of a time-limited lecture period. As educators, it is important to consider diverse student needs in the context of a demanding curriculum and provide available resources and a supportive academic environment that allows for individual success.

References


Elizabeth R. Whitney, Ph.D
Department of Anatomy and Neurobiology
Boston University School of Medicine
715 Albany Street
Boston, MA 02118
ewhitney@bu.edu

Monica A. Pessina, Ph.D
Department of Anatomy and Neurobiology
Boston University School of Medicine