Innovations Through Information Technology

2004 Information Resources Management Association
International Conference
New Orleans, Louisiana, USA

May 23-26, 2004

Mehdi Khosrow-Pour
Information Resources Management Association, USA

IDEA GROUP PUBLISHING

Hershey • London • Melbourne • Singapore
http://www.idea-group.com

Neuro-fuzzy Modeling for Inferring a User's Interest in a Web Page: An Empirical Study

Azene Zenebe
Department of Information Systems, University of Maryland (UMBC), 1000 Hilltop Circle, Baltimore, MD 21250, USA,
azenez1@umbc.edu

Aryya Gangopadhyay
Department of Information Systems, University of Maryland (UMBC), 1000 Hilltop Circle, Baltimore, MD 21250, USA,
gangopad@umbc.edu

Anthony F. Norcio
Department of Information Systems, University of Maryland (UMBC), 1000 Hilltop Circle, Baltimore, MD 21250, USA,
norcio@umbc.edu

ABSTRACT

Inferring about a user's interest in a Web Page and its related pages can be considered as a part of learning or constructing a user model. Application of the standard machine learning techniques for user modeling seems appropriate and effective but they have drawbacks. The need for large labeled data, 'concept drift' concerns due to the dynamic nature of user's features, and computational complexity are the three major problems. Further, the standard machine learning techniques do not handle adequately the inherent incomplete, imprecise and uncertain knowledge about users. The objective of the study is for handling of uncertainty due to imprecision and vagueness in inferring a user's interest.. This study compares the effectiveness of the Neuro-fuzzy modeling approach to artificial neural network, decision tree induction, and the Naïve Bayes. The results indicate that the Neuro-fuzzy modeling approach has potential and practical application in handling uncertainty and predicting a user interest in a Web Page and related pages.

1. INTRODUCTION

Inferring a user's interest in a Web Page and its related pages can be viewed as learning a user's behavioral characteristics which becomes part of the user model. Although the application of the standard machine learning approaches for user modeling seems appropriate and effective, they have drawbacks with respect to their requirements for large labeled data, 'concept drift' concerns due to the dynamic nature of users' feature, and computational complexity [1]. Furthermore, the following additional requirements are lacking in the standard machine learning approaches:

- handling adequately the incomplete, imprecise, uncertain and conflicting knowledge about users [2].
- learned models are required to be comprehensible to humans in order to understand, explain and maintain [3].
- considering prior knowledge to improve the learning process [4].
- effective techniques for representing semantic rich knowledge about users are required [2, 5].

For handling uncertainty due to imprecision and vagueness in inferring a user's behavioral characteristics, an empirical study is conducted to compare the effectiveness of the Neuro-Fuzzy modeling approach with Artificial Neural Network, decision tree induction, and the Naïve Bayes. The results of the study suggest that the Neuro-fuzzy modeling approach has potential and practical applications in handling uncertainty due to vagueness in addition to predicting a user interest to a Web Page and related pages. Moreover, with high granularity level of the decision variable, the Neuro-Fuzzy modeling approach is able to handle the uncertainty in the user data.

The different machine learning approaches and related work are reviewed in section two of this paper. Section three presents the research questions, research design and dataset. The results and discussion of the study are presented in section four. Finally, the conclusions, limitations and future work are presented in section five.

2. MACHINE LEARNING APPROACHES AND RELATED WORK

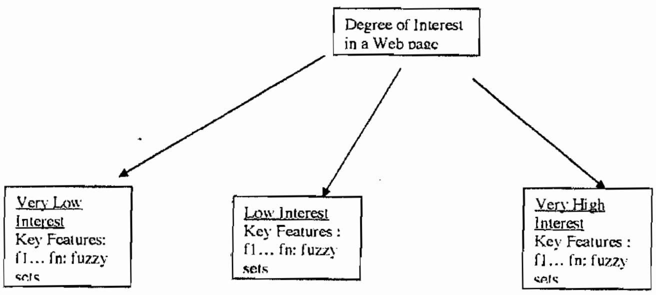
For the purpose of this study learning approaches are categorized into two categories: standard learning approaches and soft computing based learning approaches referred as soft machine learning. From the former category improved decision tree induction (ID3) algorithm C4.5/J48 and the Naive Bayes classifiers [6] are examples. From the latter category Artificial Neural Network (ANN) and Neuro-Fuzzy modeling [7] are examples.

Zadeh proposes [8] the soft computing paradigm that includes fuzzy set theory, genetic algorithms, and neural networks. It differs from classical (hard) computing paradigm in that the soft computing paradigm is tolerant to imprecision, uncertainty, partial truth, and approximation similar to the human mind. The hybrid of neural networks and fuzzy set theory form a neuro-fuzzy modeling technique [7, 9]. Consequently, neuro-fuzzy approach offers the aforementioned advantages, where fuzzy sets provide the means to handle the uncertainty, and ANN provides the means to learn the parameters of the fuzzy system from empirical data and expert knowledge. For technical information on the representation and learning of the network, interested reader can refer to [7, 9].

Many studies have been reported on the success of neuro-fuzzy modeling in control systems, i.e., [7, 10, 11]. However, there are few studies that employ neuro-fuzzy modeling for inferring about user characteristics and constructing a user model. The work by Stathacopoulou and et al. [12] uses the neuro-fuzzy synergism to evaluate students regarding their knowledge and cognitive abilities in the physics domain area in the context of an intelligent tutoring system. Also, the application of Neuro-fuzzy modeling in other related systems such as web mining, information retrieval and data mining is emerging, e.g.,[11, 13, 14]

Figure 1 (a) and (b) show examples of the learning task and structure of the fuzzy classes. Unlike crisp classification, in this fuzzy classifica-

Figure 1: The learning task (a); and the fuzzy classes or stereotypes for Degree of Interest (b)



Where $f_1, f_2, \ldots f_n$ are user's features or attributes

tion a user's degree of interest could be in one or more fuzzy classes with varying degree of membership along with a measure of the degree of (un)certainty.

3. RESEARCH QUESTION, RESEARCH DESIGN AND DATA SET

The general research question is whether modeling a user's interest in a Web page using neuro-fuzzy approach performs more effectively than standard machine learning approaches? This question tries to addresses the following sub questions. Are the theoretical benefits of neuro-fuzzy modeling in handling uncertainty realistic? What is the effect of granularity in the learning?

In this study, X_k, for k from 1 to n, represents the n input features or factors that are useful for inferring the output features Y_m, for m from 1 to L. Inferring about degree of interest in a Web page (Y) based on user interaction data is used as a rudimentary case. Currently, the most common technique is to ask users to rate a Web page explicitly. The limitations of this explicit rating, and the need for better techniques such as inferring from user actions are presented in the literature [15, 16]. Claypool and et al. [16] have studied the relationship between various implicit actions and the explicit ratings for Web pages based on over 40 hours of Web browsing by over 70 students. They have found that the time spent on a page (X1), the amount (time) of scrolling on a page (X2), and the combination of time and scrolling have a strong positive relationship with explicit interest (Y). Hence, the data collected and made available to public by the authors [16] is used.

In addition, the descriptive statistics show that the data is highly dispersed, and hence can be considered as data with outliers and noise. The time measures used for the input variables are imprecise because it is assumed that the user spent all the time browsing that specific Web page. Moreover, the concept "user interest in a Web page" is inherently vague and imprecise. All these add to the uncertainty of the learning process and the learned model.

In order to assess the effect of uncertainty due to outliers and noise, the dataset is organized into two categories: data set with uncertainty (DSWU), and data set without or reduced uncertainty (DSWOU). The dataset without uncertainty (DSWOU) is data having explicit rating, and with noise and outliers filtered or removed. Nevertheless, in both cases, uncertainty due to the vagueness and imprecision cannot be eliminated, but it is represented using fuzzy set. Specifically, the simplified form of TSK (Takagi-Sugeno-Kang) fuzzy model where the outputs are expressed by constant values is used [9].

With respect to machine learning approaches, the two scenarios are conventional Machine Learning (CML) and Soft Machine Learning (SML). From CML the improved decision tree induction algorithm C4.5/J48 and the Naive Bayes classifier implemented in the Weka machine learning tool [6] are selected. From SML, the Artificial Neural Network (ANN) implemented in the Weka machine learning tool[6] and Neuro-Fuzzy classification tool NEFCLASS implemented by Nauck [17] are selected.

Table 1: Experimental Setup

DS[1] MLA[2]	DSWOU[3]		DSWU[4]	
	5C	3C	5C	3C
CML	C4.5, Naïve Bayes, 955[6]	C4.5, Naïve Bayes, 955	C4.5, Naïve Bayes, 1821	C4.5, Naïve Bayes, 1821
SML	ANN[5], NEFCLASS[7], 955	ANN, NEFCLASS, 955	ANN, NEFCLASS, 1821	ANN , NEFCLASS, 1821

1 DS = Data Set
2 MLT=Machine Learning Technique
3 DSWOU= Dataset Without/reduced Uncertainty
4 DSWU = Dataset With Uncertainty
5 ANN =Artificial Neural Network
6 The numbers are the total number of instances used for training
7 NEFCLASS = Neuro-Fuzzy classifier

For the purpose of assessing the effect of granularity on the performance of the learning, granularity that has semantic sound and complete coverage (some psychological findings suggest the upper limit of 7±2 linguistic terms) are considered [18]. With respect to the granularity of the class variable (Y), we consider: Five classes (5C): 1=least interesting, 2=less interesting, 3=just interesting, 4= more interesting, 5=most interesting; and three classes (3C): 1=less interesting, 2=just interesting, 3= highly interesting. Similarly, for X1 and X2 granularity 3 and 5 are considered. Finally, the ten-fold cross-validation is used for determining the accuracy (as determined as the percentage of correct classification) and ambiguity of the classification (as determined by root mean squared error (RMSE)) of each of the machine learning approaches. The experimental setup is summarized in Table 1.

After exploring the different learning parameters and options of the different learning approaches, optimal values are used. In particular, for Neuro-fuzzy technique the most successful and widely used Triangular membership function, and the five fuzzy linguistic terms are used for X1 and X2. For ANN 1 hidden layer with three nodes, learning rate 0.3 and 500 epochs, momentum 0.2 and normalized attribute are used.

4. RESULTS AND DISCUSSION

Table 2 presents the summary of the results, and the following lessons are learned about the effectiveness of the neuro-fuzzy approach compared to the other approaches.

i) The overall performance of neuro-fuzzy approach is greater than or equal to the performance of others. It is also better than random classification.

ii) With high granularity neuro-fuzzy approach and ANN tolerate the uncertainty in user data. Whereas the performance of CML decreases with uncertainty in the user data.

iii) With low granularity, Neuro-fuzzy approach and ANN still tolerate the uncertainty in the user data, but performance

Table 2: Summary of the results

Data Set	DSWOU		DSWU	
MLA*	Accuracy/ RMSE[1]		Accuracy/ RMSE	
	5C	3C	5C	3C
CML				
•C4.5	27.11/0.41	48.80/0.43	26.23/0.41	45.10/0.46
•Bayesian			24.78/0.41	42.10/0.46
SML				
•ANN	25.11/0.39	49.00/0.44	26.24/0.40	45.08/0.46
•NEFCLASS	29.00/0.7	49.00/0.5	28.17/0.72	45.00/0.55

1 … RSME = Root Mean Square Error

decreases (for ANN by 3.82, for NEFCLASS by 3.65) to lesser extent compared to the case of high granularity. Whereas, the performance of CML decreases to greater extent than Neuro-fuzzy approach and ANN (for C4.5 by 3.70, for Bayes by 11.74) when the user data is with uncertainty.

iv) The complexity of the model built using Neuro-fuzzy approach (consists of the learned fuzzy sets and fuzzy rules) is simple and easy to be comprehended by human compared to the model obtained by ANN and other CML approach.

5. CONCLUSION, LIMITATION AND FUTURE WORK

This study tries to ascertain the usefulness of the neuro-fuzzy approach for inferring about a user. In conclusion, not only is the neuro-fuzzy approach handles uncertainty well but also it performs well in inferring user interest in a Web page. With high granularity level of the decision variable, the neuro-fuzzy and ANN techniques are able to handle the uncertainty due to the noise and outliers in the user data. Some limitations in this study are:

i) the comparison is based only on a single dataset,
ii) other attributes like book marking, number of visits and user domain knowledge need to be incorporated in the user model.
iii) inclusion of prior knowledge, and assessing the effect will be considered in the future study, and
iv) measuring the uncertainty of the inference process is also left for future study.

REFERENCES

[1] G. I. Webb, M. J. Pazzani, and D. Billsus, "Machine Learning for User Modeling," *User modeling and user-adapted interaction*, vol. 11, pp. 19-29, 2001.

[2] Q. Chen and A. F. Norcio, "Knowledge Engineering in Adaptive Interface and User Modeling.," in *Human Computer Interaction: Issues and Challenges*, Q. Chen, Ed. PA: Idea Group Publishing, 2001, pp. 113-133.

[3] C. Chiu, A. F. Norcio, and C.-I. HSU, "Reasoning on domain knowledge level in Human-Computer Interaction," *Information Sciences*, vol. 1, pp. 31-46, 1994.

[4] D. Nauck and R. Kruse, "New learning strategies for NEFCLASS," presented at Seventh International Fuzzy Systems Association World Congress IFSA'97, Prague, 1997.

[5] Q. Chen and A. F. Norcio, "Modeling a user's domain knowledge with neural networks," *International Journal of Human-Computer Interaction*, vol. 9, pp. 25-40, 1997.

[6] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. CA: The Morgan Kaufmann, 2000.

[7] C.-T. Lin and C. S. G. Lee, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*. NJ: Prentice Hall, 1996.

[8] L. A. Zadeh, "Fuzzy Sets," *Information Control*, vol. 8, pp. 338-353, 1965.

[9] D. Nauck, "Neuro-Fuzzy Systems: Review and Prospects," presented at Fifth European Congress on Intelligent Techniques and Soft Computing (EUFIT'97), Aachem, 1997.

[10] K. Sbujace, S. Sorathy, R. Nicholson, and R. Getorge, "Neuro-fuzzy Controller And Convention Controller: A Comparison," presented at Proceedings of the 5th Biannual, 2002.

[11] A. Abraham, "Abraham A. Neuro-Fuzzy Systems: State-of-the-Art Modeling Techniques," presented at The Sixth International Work Conference on Artificial and Natural Neural Networks, IWANN 2001, 2001.

[12] R. Stathacopoulou, G. D. Magoulas, and M. Grigoriadou, " Neural network-based fuzzy modeling of the student in intelligent tutoring systems," presented at Proceedings of the INNS-IEEE International Joint Conference on Neural Networks, Washington, U.S.A., 1999.

[13] S. Mitra, S. K. Pal, and P. Mitra, "Data Mining in Soft Computing Framework: A Survey," *IEEE Transactions on Neural Networks*, vol. 13, pp. 3-14, 2002.

[14] S. K. Pal, V. Talwar, and P. Mitra, "Web mining in soft computing framework: relevance, state of the art and future directions," *IEEE Transactions on Neural Networks*, vol. 13, pp. 1163-1177, 2002.

[15] J. B. Schafer, K. J. and J. Riedl, "Electronic Commerce Recommender Applications.," *Journal of Data Mining and Knowledge Discovery*, vol. 5, pp. 115-152, 2001.

[16] M. Claypool, P. Le, M. Waseda, and D. Brown, " Inferring User Interest," *IEEE Internet Computing*, vol. 5, pp. 32 - 39, 2001.

[17] D. Nauck, "Design and Implementation of a Neuro-Fuzzy Data Analysis Tool in Java," in *Institute of Knowledge Processing and Language Engineering, Faculty of Computer Science.*. Magdeburg: University of Magdeburg, 1999, pp. 103.

[18] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy Sets*. Cambridge, Massachusetts: The MIT Press, 1998.