A Neuro-Fuzzy Strategy for Web Personalization

G. Castellano, A.M. Fanelli, P. Plantamura, M.A. Torsello

Computer Science Department, University of Bari
Via Orabona 4, 70126 Bari, ITALY
E-mail: {castellano, fanelli, pplantamura, torsello}@di.uniba.it

Abstract

In this paper, we propose the use of a neuro-fuzzy strategy to develop a Web personalization framework for the dynamic suggestion of URLs retained interesting for the currently connected users. In particular, a hybrid strategy exploiting the combination of the fuzzy logic with the neural paradigm is proposed in order to discover useful knowledge from session data identified from the analysis of log files and represent it in a set of fuzzy rules expressed in an interpretable form.

Introduction

The growing quantity of information and applications available on the World Wide Web has recently imposed some kind of personalization tool for the Web information space. A Web Personalization System (WPS) can be considered as a process that aims at adapting the Web information space with a user's interaction in order to provide users with the information they are interested in, without asking for it explicitly [1,2,3]. In this way, WPSs appear as intelligent and automatic processes, since adaptation to the user's interests is achieved in a transparent manner for the users connected to the site.

Generally, two different processes can be distinguished in a WPS: a knowledge discovery engine and a recommendation engine. In the knowledge discovery engine, Web usage mining techniques are employed in order to extract the knowledge about the user's preferences. In the recommendation engine, the knowledge extracted through the previous analysis is used to provide recommendations to the users, such as adding hyperlinks to the last web page requested by the user, depending on the type of user [4].

One key feature to be addressed in developing a WPS is vagueness and imprecision characterizing Web information space. Besides, it includes a great deal of uncertainty which pervades all stages of the user's interaction: from the definition of the user access model to the recommendation process. Different works have proposed Web Personalization frameworks based on the use of traditional machine learning techniques [5]. However, these techniques may be inefficient and inadequate to handle the uncertainty characterizing the Web interaction processes. These limitations can be overcome through the application of soft computing techniques (i.e. fuzzy logic, neural networks, neuro-fuzzy systems, etc…) which are able to properly deal with the imprecision, uncertainty and partial truths underlying the personalization process [6].

In this paper, we investigate the use of a neuro-fuzzy strategy to develop a Web Personalization framework. In such a framework, a hybrid approach based on the combination of the fuzzy reasoning with a neural network is employed in order to determine user profiles and to provide dynamical predictions about web pages to be suggested to the current user, according to the user profiles previously identified.

The rest of the paper is organized as follows. The next section describes the working scheme of the proposed framework. Finally, some conclusive remarks are drawn.

The proposed framework

Two principal activities can be distinguished in our framework: extraction of user profiles and recommendation of URLs. Extraction of user profiles starts from the analysis of the data included in a Web log file to identify all user sessions and ends up with a clustering of the sessions to obtain a number of user profiles expressed by the visited URLs.

More precisely, we start the personalization process from a Web log file containing the information regarding all the accesses of the users on a Web site (*access time, IP address, URL viewed,…etc*). A user session is hence represented by the accesses originating from the same IP address within a predefined time period. If the Web site is composed by n valid URLs, then each URL is assigned to a unique number $j \in 1,...,n$. In this way, the ith user session is represented by an n-dimensional binary vector $\mathbf{s}^{(i)} = \left(s_1^{(i)}, s_2^{(i)},..., s_n^{(i)} \right)$ with the property:

$$s_j^{(i)} = \begin{cases} 1 & if\ user\ accessed\ j^{th}\ URL \\ 0 & otherwise \end{cases}$$

Starting from the collection of user sessions, a standard fuzzy clustering algorithm is applied in order to categorize the Web sessions so as to obtain a number of clusters which represent user profiles.

The next main activity of our framework is recommendation of Web pages according to the extracted user profiles. In particular, for each user category, the recommendation system suggests the relevance of each

URL, i.e. provides an n-dimensional vector defined as $\mathbf{r}^{(i)} = \left(r_1^{(i)}, r_2^{(i)}, ..., r_n^{(i)} \right)$, where $r_j^{(i)}$ represents the relevance of the jth URL for the ith profile. Thus, the recommendation problem can be stated as that to derive a model that predicts relevance values r_j given a new user session $\mathbf{s}$. The model considered in this work is represented by a fuzzy rule base where each rule is expressed in the form:

IF (s_1 **is** A_{1k}) **AND ... AND** (s_n **is** A_{nk})
THEN (relevance of URL$_1$ **is** r_{1k}) **AND ...**
AND (relevance of URL$_n$ **is** r_{nk})

Here, $k = 1,...,K$ where K is the number of fuzzy rules, A_{ik} ($i = 1,...,n$) are fuzzy sets defined over the input variables s_i and r_{jk} ($j = 1,..., n$) are fuzzy singletons expressing the relevance of the jth URL suggested for the current user session.

In order to automatically derive a proper number of fuzzy rules and the parameters involved in the fuzzy rules, the fuzzy model is implemented using a particular neural network so that the structure and the parameters of the fuzzy rule base can be obtained by adapting the topology and the parameters of the correspondent neuro-fuzzy network through learning from data. We adopt a hybrid learning strategy which is characterized by two learning steps. In the first learning step, the structure and the parameters of the fuzzy rules are identified through the application of an unsupervised learning process. Successively, the parameters are tuned by a supervised learning process in order to improve the accuracy of the model. Major information about the adopted neuro-fuzzy strategy can be found in [7].

Conclusive remarks

The developed neuro-fuzzy recommendation system has been applied on both synthetic and real Web usage data.
Preliminary results show the effectiveness of the system in finding good associations (in terms of precision and recall) between user sessions and dynamically recommended Web hyperlinks according to the identified user profiles.
Fuzzy clustering revealed to be an effective tool in the task of extraction of user profiles. In fact, this approach allows to derive overlapping clusters so that a user may belong to more than one category. Moreover, this is able to handle with the uncertain and fuzzy nature inherent the browsing behavior of users that change their interests over time.
The use of the neuro-fuzzy strategy for the discovery of the fuzzy recommendation rules permits to exploit the well known of the learning capabilities of neural networks with the transparent knowledge representation of fuzzy systems.

In this way, the disadvantage of the black-box behavior of the neural networks is overcome by the possibility to obtain a recommendation model expressed in a more interpretable form. In effect, each recommendation rule can be expressed in a linguistic fashion, by associating a linguistic label (LOW or HIGH) to each fuzzy set. The interpretability of the discovered recommendation model is a property particularly required for the generation of intelligent suggestions, making the neuro-fuzzy strategy suitable to the task of discovery recommendation model.

References

[1] O. Nasraoui, World Wide Web Personalization. Encyclopedia of Data Mining and Data Warehousing, J. Wang, Ed, Idea Group, 2005.

[2] M. Eirinaki and M. Vazirgiannis. Web mining for web personalization. ACM TOIT., 3(1):2–27, Feb. 2003.

[3] M. Mulvenna, S. Anand, and A. Buchner. Personalization on the net using web mining. CACM, 43(8):123–125, Aug. 2000.

[4] B. S. Suryavanshi, N. Shiri, S. P. Mudur. A Fuzzy Hybrid Collaborative Filtering Technique for Web Personalization. Proc. of 3rd Workshop on Intelligent Techniques for Web Personalisation (ITWP'05), in conjunction with the 19th International Joint Conference on Artificial Intelligence (IJCAI05), Edinburg, Scotland, UK, Aug. 2005.

[5] D. Pierrakos, G. Paliouras, C. Papatheodorou, and C. D. Spyropoulos. Web usage mining as a tool for personalization: A survey. User Modeling and User-Adapted Interaction, 13(4), 311–372, 2003.

[6] E. Frias-Martinez, G. Magoulas, S. Chen, R. Macredie. Modeling human behavior in user-adaptive systems: Recent advances using soft computing techniques. Expert Systems with Applications, 29, 320–329, 2005.

[7] G. Castellano, C. Castiello, A.M. Fanelli, C. Mencar. Discovering prediction rules by a neuro-fuzzy modeling framework. V. Palade, R. J. Howlett, L. Jain eds., Lecture Notes in Artificial Intelligence - Knowledge-Based Intelligent Information Engineering Systems, 2773, 1242-1248, Springer Verlag, Berlin Heidelberg, 2003.