Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Cognitive styles and Web-based instruction: Field Dependent/Independent vs. Holist/Serialist

Natalie Clewley, Sherry Y. Chen and Xiaohui Liu
Department of Information Systems, Mathematics and Computing
Brunel University
Uxbridge, UK
Natalie.Clewley, Sherry.Chen and Xiaohui.Liu @brunel.ac.uk

***Abstract*— As Web-based instruction (WBI) becomes increasingly popular, designers are faced with the challenge of identifying the varying preferences of learners and accommodating them in WBI programs. Cognitive style has been shown to have a significant effect on users' preferences for WBI. In particular, Witkin's Field Dependent/Independent has been widely studied in this area. It has been suggested that this cognitive style has conceptual links with the rarely studied Pask's Holist/Serialist dimension. Therefore, this study investigates the relationship between these two cognitive style dimensions by using a data mining approach that integrates feature selection and decision trees. The results show that in general Field Dependent users and Holists show the same preferences for WBI, as do Field Independent users and Serialists. However, this study also highlights the similarities in preference between Field Dependent users and Serialists, and Field Independent users and Holists. The findings of this study contribute towards the understanding of Field Dependent/Independent and Holist/Serialist users' preferences. Additionally, a novel data mining methodology is proposed.**

***Keywords*— cognitive style, user preferences, web-based instruction, data mining, feature selection, decision trees**

I. INTRODUCTION

Web-based instruction (WBI) is becoming increasingly popular because it is accessible for a wide range of learners. Such popularity is a prominent challenge to designers because they need to identify the different preferences of learners and accommodate them in WBI programs. The preferences of the different learners are dependent upon their backgrounds, skills and knowledge. Thus, human factors, the individual characteristics that affect the design of human-computer interaction [1], play an important role in the development of WBI. Among various human factors, previous studies indicate that cognitive style, an individual's approach to organizing and representing information [2], has a significant effect on users' preferences. Witkin's [3] Field Independence/Dependence is one of the most widely studied dimensions of cognitive style. In this cognitive style, users are categorized into one of three groups: Field Dependent (FD), Intermediate (I) or Field Independent (FI). The main difference between FD and FI users is that the former generally perceives objects as a whole, whereas the latter focuses more on the individual parts of the object. The differences between these two groups and their implications for the design of WBI have been well-documented. For example, Field Independent learners prefer to use the alphabetical index whilst Field Dependent learners prefer the hierarchical map [4].

Another dimension of cognitive style, Pask's Holist/Serialist [5] has conceptual links with Field Independence/Dependence. For instance, Holists prefer to take a global approach, which is similar to the Field Dependent strategy of perceiving objects as a whole. On the other hand, both Serialists and Field Independent users prefer to take a local analytical approach. Although these two dimensions of cognitive style share some similarities, they are measured with different instruments. The CSA [6] or GEFT [7] are instruments that can be used to measure Field Independence/Dependence. In contrast, the SPQ [8] is the instrument used to classify Holist/Serialist. It seems that the CSA or GEFT are more recognized and available than the SPQ. Furthermore, the reliability of the CSA and GEFT are already established [9][10], but the reliability of the SPQ is still not identified. Due to such reasons, Field Independence/ Dependence was widely investigated in previous research e.g. [11], but few studies paid attention to Holist/Serialist. To this end, the study presented in this paper examines these two dimensions of cognitive styles. More specifically, we try to answer the following research question: how do these two dimensions of cognitive style interact with each other in learners' preferences for the use of WBI? The answer to this question is significant because it will not only increase the awareness of Holist/Serialist but illustrate the relationships between these two dimensions of cognitive style. Such knowledge can help designers understand how to develop WBI programs to meet the needs of Holist/Serialists as well as effectively implement WBI programs for one dimension (e.g., Field Dependence/Independence) with design solutions for the other dimension (e.g., Holists/Serialists).

The other significance of this paper lies within the techniques used for data analyses. Previous studies analyzing the relationship between cognitive style and user preferences have used statistical techniques, such as ANOVA or factor analysis. These techniques are useful but they lack methods to control the quality of the data. In other words, the results may not be the most accurate when the data quality is low. In

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

particular, human preferences data is by nature very 'noisy', meaning it will often include data items that can be considered as outliers. Consequently, the inclusion of these data items in the statistical analyses can introduce bias in the results achieved through such techniques. Therefore, a method that is able to filter out the noisy data whilst providing the same level of analysis is needed.

Feature selection is one technique that can address this issue. Generally, feature selection techniques can be divided into two categories: filters, which produce a ranking of all features without involving any classifiers; and wrappers, which use classifiers to evaluate subsets and interactions of features [12]. Although wrappers usually provide better performance [13], traditionally wrapper methods consider just one classifier. The problem with this method is that each classifier will have its own biases. Thus, each classifier will select a different feature subset which may lead to varying levels of accuracy. Therefore, there is a need to consider multiple classifiers, instead of just the one classifier as is traditionally used. Identifying the features commonly selected by several classifiers could maximize the overall effectiveness of feature selection by making sure that only the most relevant subset of features is chosen. Once this relevant subset of features is chosen, a classification technique can be used to illustrate the relationships between the relevant features and a particular target variable (e.g., cognitive styles). Among various classification techniques, decision trees can be used to identify the accuracy of the relevant feature sets. By doing so, a decision tree with the highest accuracy can be used to demonstrate the relationships among features. The advantage of this approach is that more reliable relationships can be identified with both common relevant features and a highly accurate decision tree.

Due to such an advantage, we propose to analyze the aforementioned research question with this data mining approach that integrates feature selection and decision trees. The paper is organized as follows. Section 2 describes the methodology used to conduct the experiment and the techniques used to analyze the data. Subsequently, Section 3 presents the results of the experiment and the implications of cognitive styles on the design of WBI are also discussed. Finally, conclusions are drawn and possibilities for future work are identified in Section 4.

II. METHODOLOGICAL DESIGN

A. *Research Design*

This study involved 65 postgraduate students from a UK university. The sample was evenly divided between the genders (male = 32, female = 33) and all volunteers had the basic computing and Internet skills needed to use the WBI programs involved in the study. Participants took part in a three-phase empirical study (Figure 1).

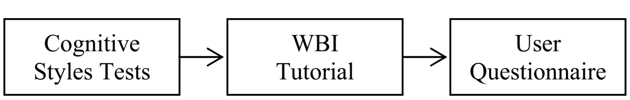

Figure 1: Research Design

All participants took part in two tests: the Cognitive Styles Analysis (CSA) and the Study Preference Questionnaire (SPQ). The former of these two tests classified participants into the Field Independence/Dependence cognitive style, whereas the SPQ identified whether participants were Holists or Serialists. The CSA is made up of two sub-tests. The first sub-test presents items containing pairs of complex geometrical figures that the individual is required to judge as either the same or different. The second sub-test involves presenting the subject with several items, each comprising of a simple geometrical shape, such as a square or a triangle, and a complex geometrical figure. The subject is then asked whether or not the simple shape is contained in the complex one by pressing one of the two marked response keys [14]. The first sub-test requires a Field Dependent capacity, whereas the second requires the disembedding capacity of Field Independence. Through these two tests, the Field Dependence competence is positively measured rather than being inferred from poor Field Independent capability [15]. This study follows the recommendations of [16], where scores below 1.03 denote Field Dependent individuals; scores of 1.36 and above denote Field Independent individuals; and scores between 1.03 and 1.35 are classified as Intermediate. The latter of these two tests consists of an 18-item inventory that assesses the strategies of Holists or Serialists. The SPQ presents subjects with two sets of statements on alternate sides and then asks them to indicate their degree of agreement with either statement or to indicate no preference [17]. This study will identify Holists and Serialists by using the following criteria: (a) if users agree with over half of statements related to Holists, they are treated as Holists; and (b) if users agree with over half of statements related to Serialists, they are treated as Serialists.

In summary, the CSA classifies users into three categories: Field Dependent, Intermediate and Field Independent; whereas the SPQ classifies users into two: Holists or Serialists. Thus, there is an imbalance in the number of categories involved in the comparison. It is therefore necessary to reconsider the CSA Intermediate category of users. This issue was resolved by calculating the mean score of the Intermediate users. Those users that scored below this mean were re-classified as Field Dependent, whilst those that scored above this mean were re-classified as Field Independent. Participant's cognitive styles were roughly evenly distributed (Table I).

After identifying cognitive style, participants were asked to interact with a WBI for approximately 90 minutes. The subject content of the WBI emphasized the practical skills of using HTML. Participants were given the freedom to explore the WBI program based on their preferences as they were provided with several navigational tools, including an alphabetical index, a hierarchical map, a main menu, section buttons and hypertext

TABLE I. COGNITIVE STYLES OF THE PARTICIPANTS

	Field Dependent	Field Independent	Total
Holist	13	20	33
Serialist	16	16	32
Total	29	36	65

links within the text. Subsequently, the participants were requested to fill out a questionnaire to identify their perceptions. The questionnaire consisted of 20 closed statements, which were designed to gather specific quantitative information about students' comprehension, preferences, and satisfaction or dissatisfaction with the WBI program, including content presentation; interaction styles; functionality and usability; and difficulties and problems. All statements used a five Likert Scale consisting of: 'strongly agree'; 'agree'; 'neutral'; 'disagree'; and 'strongly disagree'. The participants were required to indicate agreement or disagreement with each statement, by placing a check mark at the response alternative that most closely reflected their opinions.

B. Data Analysis

Data analysis used to analyze data of this study consists of two stages (Figure 2). In the first stage, feature selection was used to create subsets that are highly relevant to each of the cognitive style dimensions. These feature sets are then classified in the second stage to find the most accurate feature set. This feature set will then be used to create a decision tree that will illustrate the preferences of Field Dependent/Independent users and Holists/Serialists.

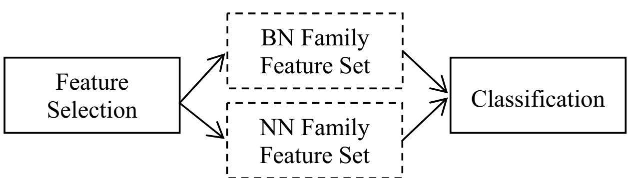

Figure 2. Data Analysis Design

1) Feature Selection

Classifiers from two different families, Bayesian Networks (BN) and K-Nearest Neighbour (KNN), were used to select relevant features. These two families were chosen because of their different biases: BN focuses on features that maximize/minimize a scoring metric whereas KNN focuses on features and instances that are deemed the 'closest' by an imposed distance metric [18]. Three classifiers from each family were chosen (Table II). These classifiers were then used to identify highly relevant feature sets, two for each of the categories of cognitive style: Field Independent, Field Dependent, Holist and Serialist. The four-step method used to attain these four feature sets for each dataset is described below.

Step 1: Each classifier algorithm was run using 10 fold cross validation. In other words, features were given a score that reflected in how many folds the feature was highlighted as relevant. For example, if the BNC classifier highlighted that Q1 was selected in 8 out of the 10 folds, the score for Q1 from the BNC would be 8.

Step 2: For each feature, the three classifier scores in each family were then averaged to give a total score of each feature

TABLE II. CLASSIFIER FAMILIES

Classifier Family	Classifier
Bayesian Network	Bayesian Networks (BNC) Naïve Bayes (NB) Averaged-One-Dependent Estimates (AODE)
k-Nearest Neighbour	Nearest Neighbour (NN) k-Nearest Neighbour (KNNC) K-Star (K*)

per family. For example, if Q1 scored 6(BNC), 6(NB) and 3(AODE), the BN Q1 score would be 5.

Step 3: The average family total was then found by averaging all the feature scores collected in Step 2.

Step 4: In the final step, any features with a Step 2 score that were higher than the Step 3 score were included in the final feature set. Thus, two feature sets for each dimension of cognitive style were collected.

2) Classification Using Decision Trees

Once these four highly relevant feature sets have been collected, the accuracy of them is verified by using decision trees. The classification is conducted using decision trees because they have been used with success in previous research in identifying characteristics of cognitive styles [e.g., [19]]. They are also easy to interpret and provide a way of measuring the accuracy of the feature sets [20]. In this way, of the two classifier family feature sets, the two that most accurately represent the preferences of Field Independent/Dependent users and Holists/Serialists, will be identified.

Among a variety of algorithms that can be used to create decision trees, C4.5 [21], CART [22] and CN2 [23] have been selected. They have been chosen because they are among the most popular, the most established and the best tested in previous research, e.g. [24]. The analysis with these algorithms will consist of three parts. Firstly, the algorithm that produces the highest average classification accuracy results will be identified. Secondly, the feature set with the highest classification accuracy will be identified from the algorithm with the highest average. Finally, this tree will be used to model the preferences of users and will be used to compare with other cognitive style groups.

III. RESULTS AND DISCUSSION

A. Feature Selection Results

As discussed in the previous section, six classifiers were used from two different families to select two relevant subsets for each cognitive style. For the Field Dependent/Independent dimension, the BN and NN classifiers selected 9 and 8 relevant features respectively (Table III). Both feature sets had five features in common:

- Q9 (*'It is hard to use back/forward buttons'*),
- Q11 *('the links provided in this tutorial help me to discover relationships between different topics')*,

- Q14 ('*After using this system I can easily use my knowledge to design home pages*'),
- Q15 ('*I found it hard to select relevant information using the map'*),
- Q19 ('*This tutorial can be used sufficiently well without any instructions'*).

The BN and NB classifiers selected 6 and 10 features relevant to Holist/Serialist (Table III). Again, both of these feature sets had five features in common:

- Q2 ('*Examples given in this tutorial are not practical*'),
- Q7 ('*I would like to have had more examples*'),
- Q9 ('*It is hard to use back/forward buttons*'),
- Q15 ('*I found it hard to select relevant information using the map*'),
- Q18 ('*It is easy to find specific information for a task with the index*').

If these findings are compared, two features appear as both relevant to Field Dependent/Independent and Holist/Serialist: Q9 and Q15. This suggests that these two features are the most relevant to each of the cognitive style dimensions. It is worth noting that both of these features refer to the way a user prefers to navigate through the subject. It implies that these two dimensions of cognitive styles have a close relationship with users' navigation preferences.

B. Classification Results

Once the feature sets were identified, classification was then performed to identify the most accurate feature set. This feature set was then used to build a decision tree, which can illustrate users' preferences of WBI programs.

To perform the classification, three different algorithms were used to build decision trees. Their classification accuracies were then calculated. Table IV shows the classification accuracies for the Field Dependent/Independent feature sets. The C4.5 algorithm had the best overall average and performed the most accurately using the KNN feature set. Table V shows the classification accuracies for the Holist/Serialist feature sets. The CN2 algorithm performed the most accurately, with both feature sets having the same highest classification accuracies. Having identified the feature sets that produced the highest classification accuracies, they are assumed to be the ones that most accurately contain the characteristics of each cognitive style group. Therefore, these feature sets are presumed to most accurately represent the different types of users' preferences for WBI. Figure 3 shows the decision tree produced for the most accurate Field Dependent/Independent feature set. Figures 4 and 5 show the decision trees produced from the two equally accurate Holist/Serialist feature sets. When these chosen decision trees are compared, three main similarities can be seen through features Q9, Q6 and Q18. Q9 ('It is hard to use back/forward buttons') appears prominently in both the FI/FD and the Holist/Serialist trees. In both, the differences between the two categories are clear. For example, in the FI/FD tree, FI users are shown to strongly disagree with Q9 whereas FD users agree or strongly agree. Alternatively, in the Holist/Serialist tree, Holists are shown to agree whereas Serialists are shown to disagree/strongly disagree. Back/forward buttons are used to help users explore the content in a non-linear way. This suggests that FI users and Serialists may feel more comfortable with non-linear navigation. Conversely, non-linear navigation may be difficult for FD users and Holists. This feature demonstrates the relationship between FI users and Serialists and between FD users and Holists.

TABLE III. NUMBER OF FEATURES SELECTED PER CLASSIFIER

Cognitive Style	Feature Set	Number of Features Selected
Field Dependent/ Independent	BN Family	9
	NN Family	8
Holist/Serialist	BN Family	6
	NN Family	10

TABLE IV. FIELD DEPENDENT/INDEPENDENT CLASSIFICATION ACCURACY (%)

Feature Set	Decision Tree Algorithm		
	C4.5	*CART*	*CN2*
BN	**87.6923**	81.3218	90.76923
KNN	**95.3846**	81.3218	86.15385
Total Average	**91.53845**	81.3218	88.46154

TABLE V. HOLIST/SERIALIST CLASSIFICATION ACCURACY

Feature Set	Decision Tree Algorithm		
	C4.5	*CART*	***CN2***
BN	72.3077	79.79925	**80**
KNN	73.8462	76.79925	**80**
Total Average	73.07695	78.29925	**80**

Q6 ('*I would have found it more helpful to be given a suggested route through this tutorial*') also appears in both trees. The FI/FD tree shows that FI users disagree with this question whereas FD users agree or strongly agree. The Holist/Serialist tree shows that Holists strongly disagree with this question and Serialists strongly agree. Unlike the results of Q9, this indicates that FI users and Holists have similar preferences whereas FD users are like Serialists. More specifically, the former can explore the content on their own while the latter need more guidance in finding their way around a topic.

Q18 ('*It is easy to find specific information for a task with the index*') is another feature selected by both trees. Both FI users and Serialists strongly agree or agree with this statement, whilst FD users and Holists disagree. The index used in this study lists all of topics in an alphabetical order so that learners can easily locate specific information. This suggests that such a mechanism is useful to FI users and Serialists but it may not be suitable for FD users and Holists. This may be due to the fact that FD users and Holists are interested in a global picture of the content, instead of a specific item.

In addition to these similarities, there are three other features that appear on the Holist/Serialist tree. Q11 (‘*the links provided in this tutorial help me to discover relationships between different topics*’) is selected in the Holist/Serialist tree, with Holists strongly disagreeing to this statement and Serialists strongly agreeing to it. It is in the nature of Holists to jump to different topics [25] so links are helpful for them. On the other hand, Serialists tend to study topics sequentially so there may be no need for them to use links. Although this feature does not appear in the chosen Field Independence/Dependence tree, it is shown in the tree that was drawn using CN2 and the NN family feature set. It can be seen in this tree that FD users agree with Q11. This suggests that FD users may have a similar preference to Holists for this question.

Q2 (‘*Examples given in this tutorial are not practical*’) and Q7 (‘*I would like to have had more examples*’) were also selected in the Holist/Serialist tree but not in the Field Independence/Dependence tree. Examples are another way of presenting content so this suggests that the content presentation may be more relevant to Holists and Serialists than to Field Dependent and Field Independent users.

IV. CONCLUSION

The relationship between two dimensions, Field Independence/Dependence and Holist/Serialist, was examined in this study through analyzing users’ preferences for WBI programs. To enable the interactions between features to be properly examined, this study proposed a data mining approach which integrates feature selection and decision trees. The former is applied to select features particularly related to cognitive styles whereas the latter visually illustrates the relationships between cognitive styles and user preferences.

The contributions of this study are threefold. Firstly, this study helps to deepen the understanding of the differences between Field Dependent and Field Independent users, and Holists and Serialists. Nevertheless, only a small sample was used in this study. Although these preliminary results helped to identify the value of continuing research in this area, there is a need to conduct further studies with a larger number of participants. This will also be particularly useful for our data mining approach.

Secondly, it has increased the understanding of the relationship between the two cognitive style dimensions. The findings of this study provide empirical evidence that in general Field Independent users and Serialists have similar preferences. Conversely, Field Dependent users are alike to Holists. However, contradictions to these rules are found in Q6 (‘*I would have found it more helpful to be given a suggested route through this tutorial*’). Consequently, previous literature identifying the differences of Field Dependent and Field Independent users can perhaps be used to suggest the differences between Holists and Serialists, but bearing in mind

Decision Tree Key

SA	= Strongly Agree	**FI**	= Field Independent
A	= Agree	**FD**	= Field Dependent
N	= No Preference	**H**	= Holist
D	= Disagree	**S**	= Serialist
SD	= Strongly Disagree		

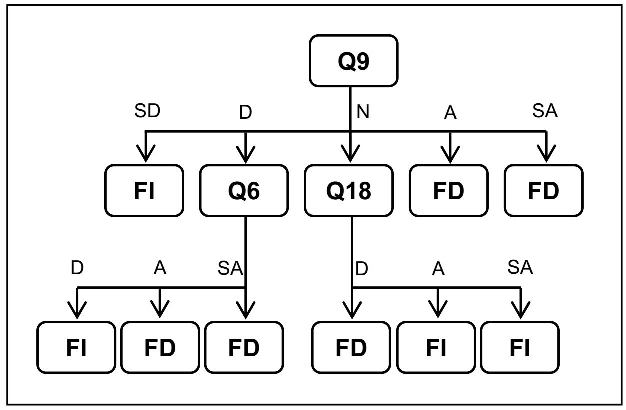

Figure 3. Field Independence/Dependence Decision Tree

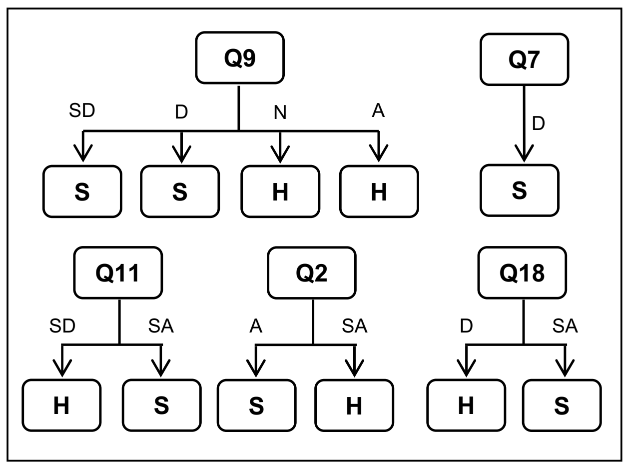

Figure 4. Holist/Serialist Decision Tree 1

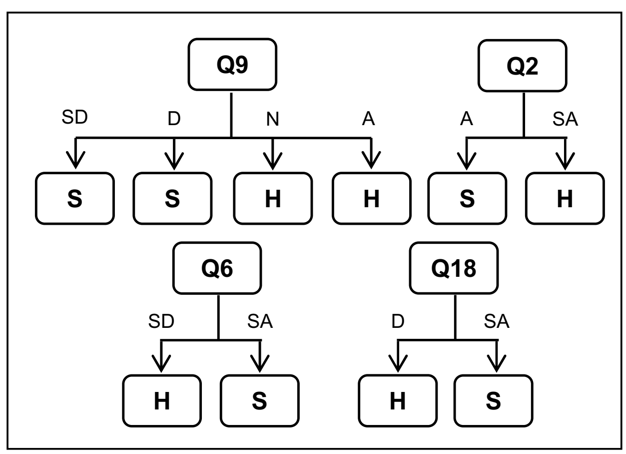

Figure 5. Holist/Serialist Decision Tree 2

the contradictions to the relationship. Note that these two dimensions of cognitive style were only identified by using the CSA and SPQ, and further work should consider using other instruments to categorize cognitive styles.

Finally, this study analyzes the data using a data mining approach, integrating feature selection with decision trees. However, classifiers from only two families were used to select relevant feature and only three algorithms were used to build decision trees. It would be interesting to extend the range of classifiers and decision tree algorithms to see if similar results are obtained.

V. REFERENCES

[1] S. Treu, User Interface Design. New York: Plenum Press,1994.

[2] R. Riding, & S. G. Rayner, Cognitive Styles and Learning Strategies. London: David Fulton Publisher, 1998.

[3] H.A. Witkin, “Cognitive style in academic performance and in teacher-student relations”, In S. Messick (Ed.) Individuality in learning: Implications of cognitive style and creativity for human development. San Francisco: Jossey-Bass, 1976.

[4] S.Y. Chen, and X. Liu, “An integrated approach for modeling learning patterns of students in Web-based instruction: a cognitive style perspective”, ACM Transactions on Computer-Human Interaction, vol. 15, no. 1, 2008, pp.1-28.

[5] G. Pask, “Styles and strategies of learning”, British Journal of Educational Psychology, vol. 46, 1979, pp.128-148.

[6] R. Riding, “Cognitive styles analysis—research applications”. Birmingham: Learning and Training Technology, 1998.

[7] H. Witkin, P. Oltman, E. Raskin and S. Karp, “A manual for the embedded figures test”. Palo Alto, CA: Consulting Psycologists Press, 1967.

[8] N. Ford, “Learning styles and strategies of postgraduate students”. *British Journal of Educational Technology*, vol. 16, 1985, pp.65-79.

[9] E.R. Peterson, I.J. Deary and E.J. Austin, “The reliability of Riding’s Cognitive Style Analysis test”, Personality and Individual Differences, vol. 34, no. 5, 2003, pp.881-891.

[10] H.A. Witkin, P.K. Oltman, E. Raskin and S.A. Karp, “Group Embedded Figures Test Manual”. Palo Alto, CA: Consulting Psychologist Press, 1971.

[11] S.Y. Chen, G.D. Magoulas and D. Dimakopoulos, “A flexible interface design for web directories to accommodate different cognitive styles”, Journal of the American Society for Information Science and Tecnology, vol. 56, no. 1, 2005, pp.70-83.

[12] J.Yang & S. Olafsson, “Optimization-based feature selection with adaptive instance sampling”, Computers and Operations Research, vol. 33, no. 11, 2006, pp. 3088-3106.

[13] B. Raman, and T.R. Ioerger, “Instance based filter for feature selection”, Journal of Machine Learning Research, 2002, pp.1–23.

[14] R.J. Riding, and M. Grimley, “Cognitive style, gender and learning from multi-media”, British Journal of Educational Technology, no. 30, vol. 1, 1999, pp.43-56.

[15] N. Ford, and S.Y. Chen, “Matching/mismatching revisited: an empirical study of learning and teaching styles”, British Journal of Educational Technology, vol. 32, 2001, pp.5-22.

[16] R.J. Riding, Cognitive styles analysis. Birmingham, UK: Learning and Training Technology, 1991.

[17] N. Ford, “Learning styles and strategies of postgraduate students”, British Journal of Educational Technology, vol. 16, 1985, pp.65-79.

[18] A.J. Gammerman, Machine learning: progress and prospects. Egham, Surrey Royal Holloway, University of London, 1997.

[19] H. Liu & V. Kešelj, “Combined mining of web server logs and web contents for classifying user navigation patterns and predicting users' future requests”, Data and Knowledge Engineering, vol. 61, no. 2, 2007, pp.304-330.

[20] P. Hsu, R. Lai & C.C. Chiu, “The hybrid of association rule algorithms and genetic algorithms for tree induction: an example of predicting the student course performance”, Expert Systems with Applications, vol. 25, no. 1, 2003, pp.51-62.

[21] J.R. Quinlan, C4.5: Programs for Machine Learning. Massachusetts: Morgan Kaufmann, 1993.

[22] L. Breiman, J.H. Friedman, R.A. Olshen & C.J.Stone, Classification and Regression Trees, Wadsworth International Group/Probability Series, California, 1984.

[23] P. Clark & T. Niblett, “The CN2 induction algorithm”, Machine Learning, vol. 3, no. 4, 1989, pp.261-284.

[24] J.K. Kim, Y.H. Cho, W.J. Kim, J.R. Kim & J.H. Suh, “A personalized recommendation procedure for Internet shopping support”, Electronic Commerce Research and Applications, vol. 1, no. 3/4, 2002, pp.301-313.

[25] N. Ford, Web-based learning through educational informatics: information science meets educational computing. Hershey, PA: IGI Global, 2008.