Learning Ability Evaluation of Online Course Based On SVM

Bo Du, Jing Zhang, Yan Zhang

College of Foreign Languages, Agricultural University of Hebei, Baoding 071001, China

Dongming Li *

College of Mechanical and Electrical Engineering, Agricultural University of Hebei, Baoding 071001, China
** Corresponding Author*

Abstract
Taking online course as the research object, learning ability evaluation based on SVM was studied. The paper fully analyzed the elements of the evaluation system, used Web mining technology to collect 11 test items evaluating students' learning ability. According to the collection data, with penalty parameter c and kernel parameter g determined by the K-fold Cross Validation (K-CV), SVM evaluation model was established with the help of Libsvm-3.20 toolboxes. Meanwhile, the evaluation accurate rate of SVM was superior that of BPNN, the former reached 97.3%. Therefore, it was simple and feasible to use SVM to evaluate the learning ability.
Key words: SVM, LEARNING EVALUATION, WEB MINING, K-FOLD CROSS VALIDATION.

1. Introduction

With the further promotion and application of education information technology and the continuous enrichment of network teaching resources, network learning as an important way of learning has been accepted by more and more people [1]. In recent years, the achievements of quality course and network course construction have been obvious, but the phenomenon of heavy construction and light use has become the main factor that has become an obstacle of the quality course and the network course construction [2]. It was easier to build up the teaching resources, but it was difficult to use. There were many reasons, and the incomplete network learning evaluation mechanism was one of the important aspects. Therefore comprehensive and scientific evaluation of network learning had a very important significance for promoting the use of network teaching resources [3].

Compared to the traditional classroom learning, it was easier to collect some information of students' learning in the online course, such as students' registration information, online learning time, learning progress, learning exchange and other information [4]. The information could be used to evaluate students' learning process, such as learning methods, learning attitude, learning effect and other aspects, and according to the results of evaluation, students could gain timely feedback, which made students continuously improve in the follow-up learning process so as to achieve better learning results and play a real role in the formative evaluation [5].

In order to solve some problems about the formative evaluation in the current online learning, the paper chose New Horizon College English online teaching system used in Agricultural University of Hebei, used online data mining technology to mine

Web data in the process of learning, established the learning ability evaluation model based on SVM, and according to the evaluation results, students could gain timely feedback and improve their learning. All in all, the formative evaluation played a real role in the network learning process.

2. Index System Construction of Learning Ability Evaluation

2.1 Web mining

Web mining was an application of data mining in Web, which was a comprehensive technology. With the rapid development of Internet, there have been the abundant data resources on Web. The original data could be obtained from the network through Web mining, and then the potentially useful knowledge patterns were mined from the original data to meet people's need, so Web mining was a new development and application of data mining [6].

According to different objects of Web mining, Web mining can be divided into the following three categories: Web content mining, Web journal Mining and Web Structure Mining [7, 8]. It was shown in figure 1.

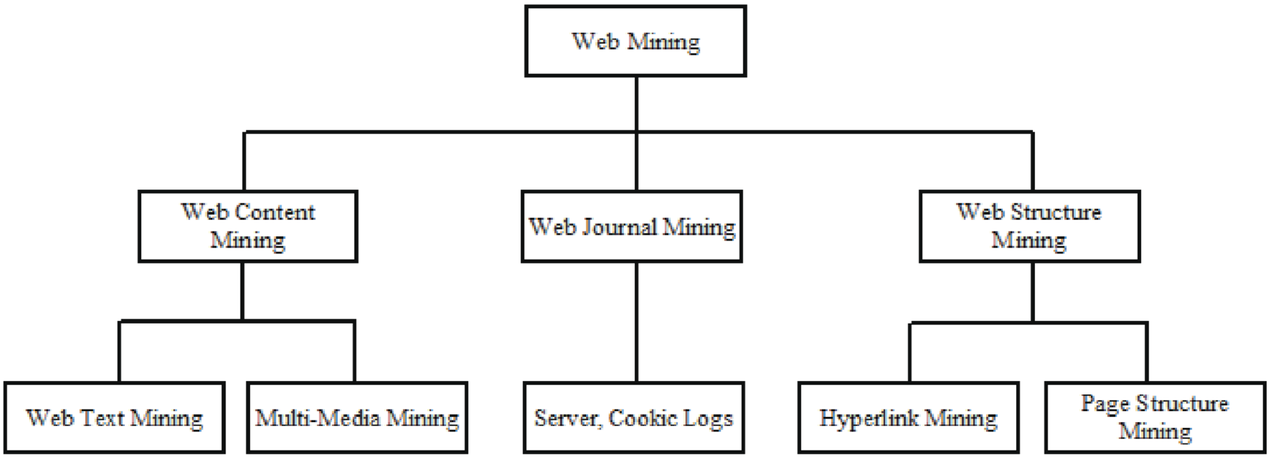

Figure 1. Categories of web mining

With the help of Web mining, all the data in the paper came from New Horizon College English online courses used in Agricultural University of Hebei.

2.2 Index System Construction of Learning Ability Evaluation

To evaluate network learning, it was critical to determine the content and evaluation criteria of network learning evaluation, that's the evaluation index system. On the basis of analyzing the existing network learning evaluation system, according to the basic principles and methods of learning evaluation, the evaluation system was constructed from 3 first-class indexes: ability of knowledge acquisition, ability of knowledge sharing and ability of knowledge absorption. There were total 11 test items under the first-class indexes. It was shown in Figure 2.

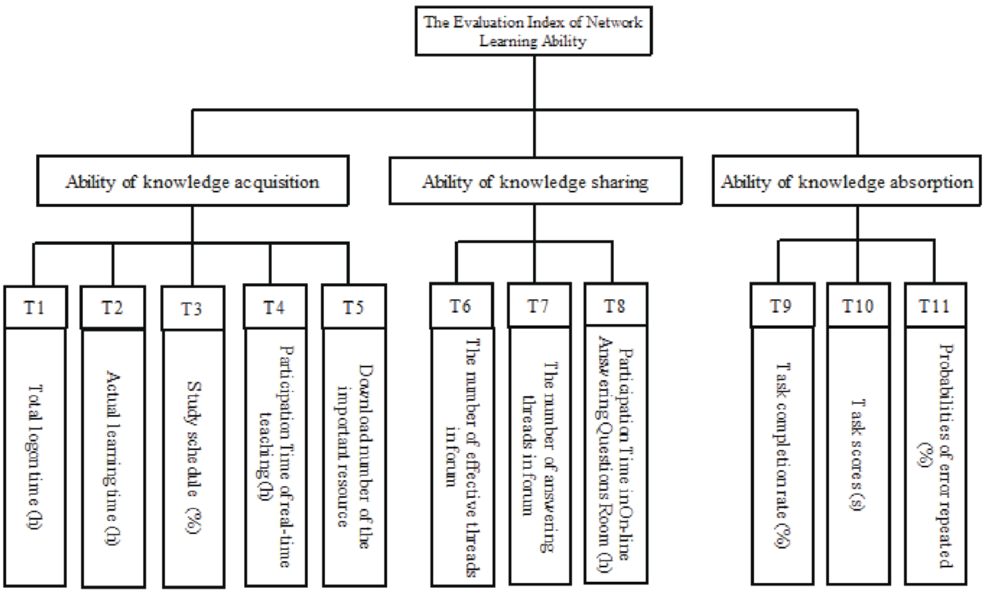

Figure 2. Evaluation index system of network learning

3. Sample Database Establishment of Learning Ability Evaluation

With English online assistant teaching system of Agricultural University of Hebei, the system has tested 314 non English major students in 10 classes of grade 2013. By Web mining and the experts' comprehensive assessment, learning ability was divided into six levels of very poor, poor, general, good, better and best, respectively written with 1, 2, 3, 4, 5 and 6.

In the 314 groups' data, group 1 to 37 learning ability was very poor, labeled "1"; group 38 to 85 learning ability was poor, labeled "2"; group 86 to153 learning ability was general, labeled "3"; group 154 to 233 learning ability was good, labeled "4"; group 234 to 275 learning ability was better, labeled "5"; group 276 to 314 learning ability was best, labeled "6". The input and output of the data was shown in Table 1, and the fractal dimension graphics was shown in Figure 3.

Table 1. Measure items data and learning ability grading

Students' number	Measure items (Input)											Ability level (Output)	Ability level labels
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11		
	20	3	30	2	0	1	0	0.5	40	60	60.6	Very Poor	1
2	20.1	3	30	2	0	1	0	0.5	40	60	60	Very Poor	1
…	…	…	…	…	…	…	…	…	…	…	…	…	…
38	25.1	4.2	40	4	6	7	2	1.3	47	67	53	Poor	2
39	25	4.1	40	4	7	7	2	1.4	47	68	55	Poor	2
…	…	…	…	…	…	…	…	…	…	…	…	…	…
86	30	7	53	8	15	16	7	4.2	60	72	45	General	3
87	31	7	55	8	15	16	6	4.4	60	73	44	General	3
…	…	…	…	…	…	…	…	…	…	…	…	…	…
154	36	9	67	13.5	22	24	10	5.5	81	74	36.8	Good	4
155	35	10	67	13.7	22	23	10	5.6	80	74	36.7	Good	4
…	…	…	…	…	…	…	…	…	…	…	…	…	…
245	55.8	19	85	21.4	27	34	16	8.9	95	82	15	Better	5
246	58	20	86	21.4	28	34	16	8.9	96	81	15.2	Better	5
…	…	…	…	…	…	…	…	…	…	…	…	…	…
313	71.9	28	100	24	24	36	25	9.7	100	97	5.1	Best	6
314	72	28	100	24	24	36	25	9.7	100	98	5	Best	6

4. Learning Ability Evaluation Model Based on SVM

4.1 SVM

Support Vector Machine (SVM), proposed by V.Vapnik etc, was a machine learning method based on statistical theory. For small sample data, it could avoid the traditional process from induction to deduction to realize high effective conduction from training sample to forecasting sample, and simplified the general classification and regression [9, 10]. The basic idea of SVM was the training data set was mapped non-linearly from the input space to Hilbert space by defining appropriate kernel function; the samples in the linear space were linearly separable to obtain classification function of optimal linear classification face in new space. The output was a linear combination of the intermediate nodes, each intermediate node corresponding to a support vector [11, 12]. As shown in figure 4.

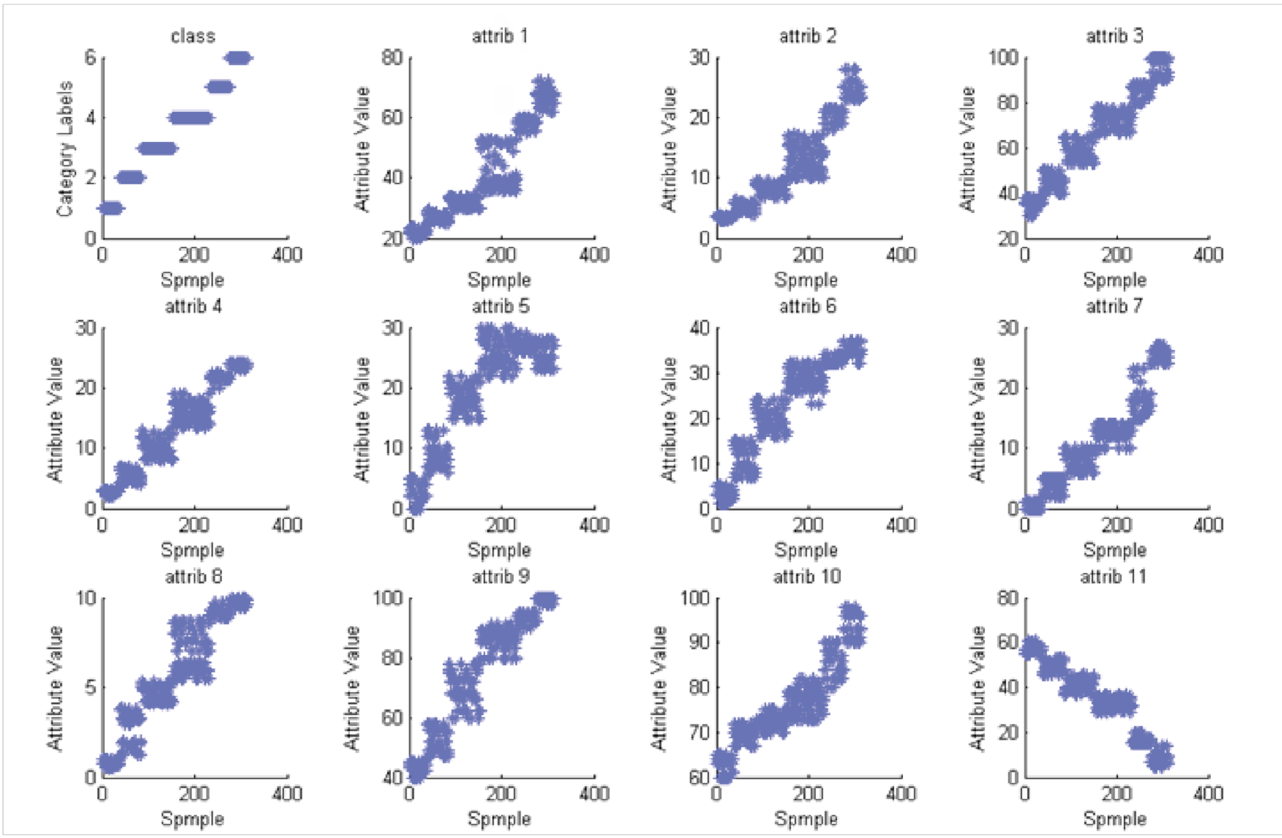

Figure 3. The fractal dimension graphics

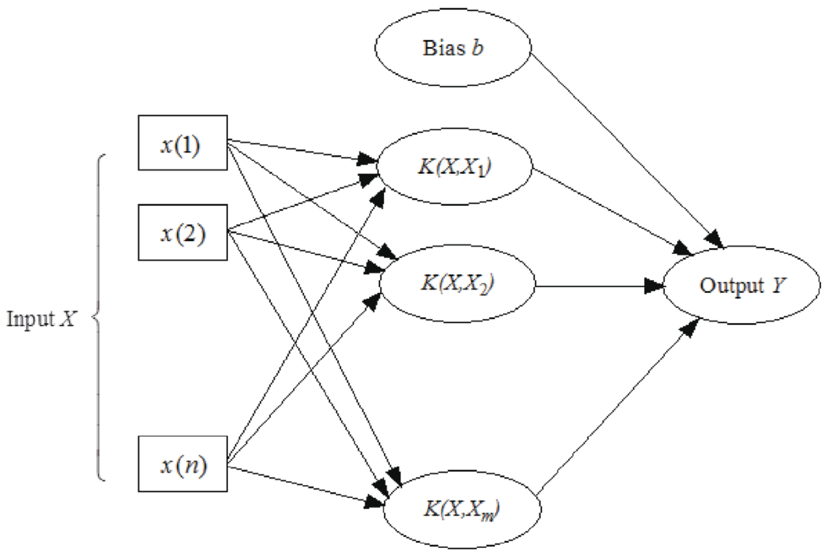

Figure 4. System structure of SVM

4.2 Evaluation Model Establishment

To establish the evaluation model, the training set and test set were extracted from the original data first, and then the sets was preprocessed (Feature extraction was needed when necessary); finally the SVM was trained by the training set to get a model which could be used to predict classification of the test set. The algorithm flow is shown in Figure 5.

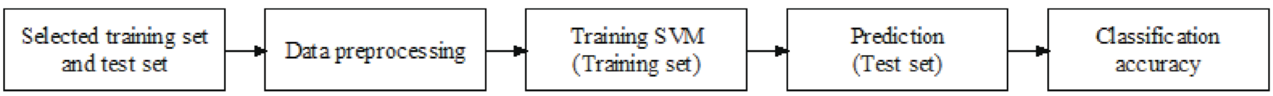

Figure 5. Algorithm flow of model

There were 2 parts in the data set that included 314 groups' data. The first part was training set having 202 groups: very poor learning ability (group 1 to 25), poor learning ability (group 38 to 70), general learning ability (group 86 to 130), good learning ability (group 154 to 200), better learning ability (group 234 to 260) and best learning ability (group 276 to 300). The second part was test set including the remaining 112 group data (group 26 to 37, group 71 to 85, group 131 to 153, group 201 to 233, group 261 to 275 and group 301 to 314).

The paper mainly used MATLAB2012 (b) software, installed Libsvm-3.20 toolbox, and made the most of its powerful computing function to realize the establishment of SVM evaluation model.

4.3 Determination of Kernel Function

There were 4 kernel functions in SVM, in which radial basis function (RBF) was used widely. RBF was a universal kernel function, which could be applied to any distribution of the sample through the choice of parameters, so in this paper the radial basis function was used.

To use the radial basis kernel function, 2 important parameters should be determined: penalty parameter *c* and kernel parameter *g*. Penalty parameter *c* controlled the balance between the empirical risk and Vapnik-Chervonenkis Dimension (VC Dimension) to achieve the compromise of the empirical risk and confidence interval [13]. The larger *c* was, the better the fitting degree of the data was. Therefore, a larger number would usually be selected to reduce the error so as to achieve the good fitting of the training samples. However, if *c* was too large, the risk of experience would increase. The kernel parameter *g* controlled the sensitivity of SVM to the change of input variables. Too large of *g* mould lower reaction of SVM; on the contrary, too small would lower anti-jamming ability [14, 15]. Therefore, it was necessary to find the optimal parameters (*c*, *g*) to achieve the optimal classification of SVM. K-fold Cross Validation was the most reliable method for small and medium scale problems. Base on this, the paper used LIBSVM3.2 software package to provide the interactive test function optimization.

Parameters *c* and *g* ranged from 2^(-10) to 2^(10), and *gstep*=2. The contour map of the selection process was shown in figure 6 and 3D View was shown in figure 7. When the parameters *c*=0.082469 and *g*=2.2974, the error is the smallest, the SVM evaluation model was the desired.

As shown in the model evaluation results, there were only 3 groups' misjudgment in 112 groups of data and accurate rate of evaluation reached 97.3 % (109/112).

In Matlab2012(b) environment, BP neural network with a 3-layer whose structure was 11-20-6 was established, with nprtool BP neural network pattern recognition toolbox adopted. After the network train-

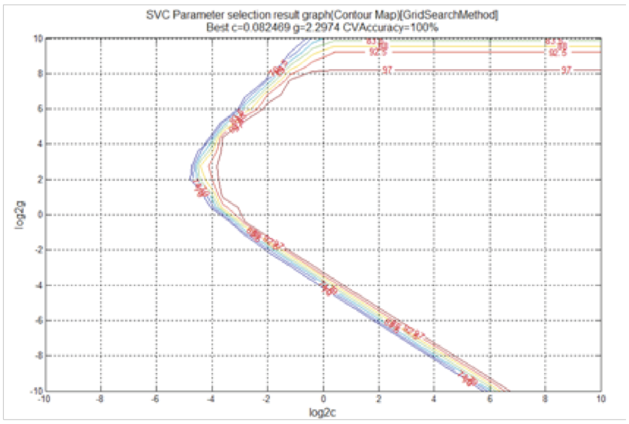

Figure 6. SVC parameter selection result graph (Contour Map)

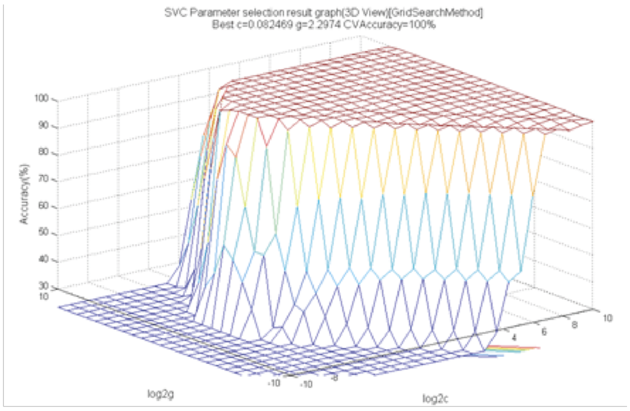

Figure 7. SVC parameter selection result graph (3D View)

5. Identification Results and Analysis

SVM trained evaluation model used, 112 test sets were evaluated in different classifications. Actual test data and the prediction data was shown in Figure 8.

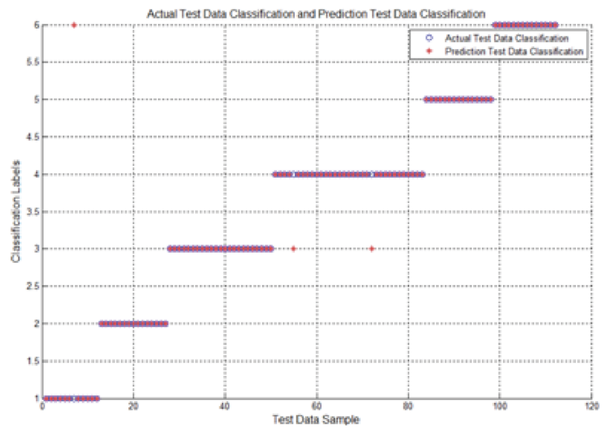

Figure 8. Actual test data classification and prediction test data classification

ing, the same test set was evaluated, and the correct recognition rate is 94.6%.

The recognition rate of SVM was slightly higher than that of BPNN, which fully reflected the superiority of SVM, the new machine learning algorithm based on statistical learning theory, for small sample statistics and learning evaluation; meanwhile, BPNN algorithm was excessively dependent on the quantity and quality of sample data in model training process.

In the case of limited training sample, the recognition rate is lower. The artificial neural network structure (the number of hidden layer neurons) also had a great effect on the recognition accuracy. If complex neural network was adopted for recognition training, the recognition rate could be improved, but the process is very complicated. All in all, SVM method can be used to evaluate the learning ability because the accuracy rate is high and it is simple and practical.

6. Conclusion

With New Horizon College English online teaching used in Agricultural University of Hebei, 11 test items reflecting students' learning ability have been mined by Web mining, and learning ability evaluation model based on SVM has been founded, whose evaluation accurate rate has reached 97.3%; overall, it has been feasible and effective to use SVM to evaluate learning ability. Furthermore, compared with the results of BP neural network algorithm, the generalization ability of SVM classifier was superior that of neural network.

References

1. Wei Xiaobao (2012) The Construction of the Evaluation System of AFLLC for China's Non-English Majors. *Contemporary Foreign Languages Studies*, No.9, p.p.61-65.
2. Dickinson, L. (1995) Autonomy and motivation: A literature review. *System*, No.2, p.p.165-174.
3. Wolters, C. (2003) Understanding procrastination from a self-regulated learning perspective. *Journal of Educational Psychology*, 95(1), p.p.179-87.
4. Littlewood, W. (1996) Autonomy: An anatomy and a framework. *System*, No.4, p.p.427-435.
5. Benson P. 2007. Autonomy in language teaching and learning. *Language Teaching* (1): 21-40.
6. Yang Man, Gou Xue-rong, Li Jian-wei (2012) Research on Learning Ability Evaluation Model Based on Network onine Learning. *Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition)*, 13(1), p.p.13-19.
7. Pearson L C & Moomaw W. (2005) The relationship between teacher autonomy and stress, work satisfaction, empowerment and professionalism. *Educational Research Quarterly*, 29(1), p.p.38-54.
8. Wang Hua-li (2010) English Learning Assessment in Higher Vocational Colleges. *Journal of Chengde Petroleum College*, 12(4), p.p77-79.
9. Cheng Hong, Shi Zhixing, Yao Wei, et al. (2010) Corn Breed Recognition Based on Support Vector Machine. *Transactions of the Chinese Society for Agricultural Machinery*, 40(4), p.p.180-183.
10. Bai Peng, Zhang Xi-bin, Zhang Bin, et al. (2008) *Support Vector Machine and its Application in Mixed Gas Infrared Spectrum Analysis*. XI AN: Xidian University Press.
11. Shi Feng, Wang Xiao-chuan, Yu Lei, et al. (2010) *MATLAB NN 30 Case Studies*. BEI JING: Beihang University Press.
12. Si Xunlian, Zhang Xufeng (2014) Comprehensive evaluation of groundwater quality based on support vector machine. *Chinese Journal of Environmental Engineering*, 8(10), p.P.4340-4344.
13. Samanta B. (2004) Gear fault detection using artificial neural network and support vector machines with genetic algorithms. *Mechanical Systems and Signal Processing*, 18(3), p.p.625-624.
14. Feng-yun Xie, San-mao Xie (2014) Gear fault classification based on support vector machine. *Hydromechatronics Engineering*, 42(18), p.p.54-57.
15. Shan Yugang, Wang Hong, Dong Shuang (2012) Improved multiclassification algorithm of one-against-one SVM. *Computer Engineering and Design*, 33(5), p.p.1837-1841.