A General Bayesian Network-Assisted Ensemble System for Context Prediction: An Emphasis on Location Prediction

Kun Chang Lee[1] and Heeryon Cho[2]

[1] Professor of MIS and WCU Professor of Creativity Science,
SKK Business School and Department of Interaction Science,
Sungkyunkwan University, Seoul 110-745, Republic of Korea
kunchanglee@gmail.com, leekc@skku.edu
[2] Department of Interaction Science,
Sungkyunkwan University, Seoul 110-745, Republic of Korea
heeryon@gmail.com, heeryon@skku.edu

Abstract. Context prediction, highlighted by accurate location prediction, has been at the heart of ubiquitous decision support systems. To improve the prediction accuracy of such systems, various methods have been proposed and tested; these include Bayesian networks, decision classifiers, and SVMs. Still, greater accuracy may be achieved when individual classifiers are integrated into an ensemble system. Meanwhile, General Bayesian Network (GBN) classifier possesses a great potential as an accurate decision support engine for context prediction. To leverage the power of both the GBN and the ensemble system, we propose a GBN-assisted ensemble system for location prediction. The proposed ensemble system uses variables extracted from Markov blanket of the GBN's class node to integrate GBN, decision tree, and SVM. The proposed system was applied to a real-world location prediction dataset, and promising results were obtained. Practical implications are discussed.

Keywords: Context Prediction, Location Prediction, Ensemble Methods, General Bayesian Network, GBN-Assisted Ensemble Classifier, ID3, C4.5, CART, SVM.

1 Introduction

Ubiquitous decision support systems [1] adapt to users' changing contexts to provide context-sensitive services that meet users' needs and preferences. Such adaptation requires the system to predict user's future context based on the current context data gathered from various sources, e.g., user's handheld devices, sensors embedded in the environment, and distributed databases. Context is defined as, "any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the integration between a user and an application, including the user and the application themselves [2]." The collected (context) data are used as training and test data for building context prediction models, and the constructed models are used to infer user's future context. The success of context prediction depends on the degree of accuracy of related predictors.

T.-h. Kim et al. (Eds.): FGIT 2010, LNCS 6485, pp. 294–303, 2010.
© Springer-Verlag Berlin Heidelberg 2010

Existing works have successfully employed Bayesian networks to predict future contexts: Patterson et al. [3] used dynamic Bayesian network to predict likely travel destinations on a city map; Hwang and Cho [4] proposed a modular Bayesian network model to infer landmarks of users from mobile log data collected through smart phones; and Kasteren and Kröse [5] used naive Bayesian and dynamic Bayesian networks to infer daily activities of elderly people performed inside their houses; Sánchez et al. used discrete hidden Markov model to automatically estimate hospital-staffs' activities [6].

Other works have employed decision tree classifiers and Support Vector Machines (SVM) for context prediction: Byun and Cheverst [7] used decision tree to infer the preferences of the user in an intelligent office environment; Lum and Lau [8] proposed a negotiation algorithm based on decision trees to handle content adaptation for mobile devices; and Matsuo et al. [9] used SVM to infer user's long-term properties such as gender, age, profession, and interests from location information; their system automatically learns patterns between users' locations and user properties.

While individual prediction models perform well on many context prediction tasks, better performance may be achieved by harnessing the power of multiple models. Ensemble-based systems, also known as multiple classifier systems, committee of classifiers, or mixture of experts, combine individual classifiers that make errors on different parts of data to enhance prediction performance. This paper investigates yet another kind of context prediction using General Bayesian Network (GBN)-based ensemble system; we investigate the power of GBN-*assisted* ensemble system on location prediction.

A GBN-assisted ensemble system is defined as a multiple classifier system that uses those variables inside the Markov blanket of a GBN's class node (or target node) to build an ensemble system. A Markov blanket of a node X consists of the direct parent of X, the direct successors of X, and all direct parents of X's direct successors [10] in a given Bayesian network. The Markov blanket of node X may be thought of as the minimal set of nodes that isolates X from the rest of the graph [11]. If a node is absent from the target node's Markov blanket, its value is completely irrelevant to the prediction [12]. Hence, the Markov blanket can be used to select the core variables that affect the class variable. In this paper, we first create a GBN to identify the variables inside the Markov blanket of GBN's class node, and then use those selected variables to create GBN-assisted ensemble system by merging GBN, decision tree, and/or SVM using voting, stacking, and grading combination strategies. Location prediction experiments are conducted using real-world data to evaluate the prediction accuracies of GBN-assisted two-classifier and three-classifier ensemble systems.

The contributions of this paper are twofold: (1) the finding that the GBN-assisted ensemble systems are comparable to the GBN-based ensemble systems; (2) what practical implications the GBN-assisted ensemble system has for location prediction. Section 2 describes three types of classifier combination strategies used for building ensemble classifiers for the location prediction experiment. Section 3 describes data, the experimental setup, and the experimental results of the ensemble performance evaluation. Practical implications of the GBN-assisted ensemble approach are discussed in Section 4. Conclusion and future research issues are given in Section 5.

2 Ensemble Methods: Voting, Stacking, and Grading

Many studies have shown that fusing a set of different classifiers (or ensemble of classifiers) with different misclassified instances (i.e., ones that do not overlap) will yield better classification performance over an individual classifier, which makes up the ensemble system, having the best performance [13]. The intuition is that if different classifiers make errors on different instances, the strategic combination of these classifiers can reduce the overall error to improve the performance of the ensemble system [14].

The success of ensemble system depends on achieving diversity among individual classifiers with respect to misclassified instances. There are four ways to achieve this diversity [14]: (1) use different training examples to train individual classifiers; (2) use different training parameters; (3) use different features to train classifier; or (4) combine entirely different type of classifiers. The first approach deals with incorporating various resampling techniques; bagging (or bootstrap aggregating) [15] and boosting [16] are two well known techniques. The second approach deals with using different parameter values such as weights, nodes, or layers (depending on the classifier to be trained) to train the individual classifier. The third approach deals with using different features to train the classifier; random subspace method [17] is one such method. Finally, the last approach deals with combining entirely different type of classifiers; an example would be combining decision trees, SVMs, and nearest neighbor classifiers.

In this paper, we focus on the last approach of combining entirely different type of classifiers to construct ensemble systems for location prediction experiment. We select three types of individual classifiers – decision trees, Bayesian classifiers, and SVM – and integrate them using three different combination strategies – voting, stacking, and grading. Note that prior to ensemble system creation, we first create GBN to identify the variables inside the Markov blanket of the class node; these variables, which parsimoniously describe the class node, are used to create GBN-assisted ensemble systems. We now briefly introduce each combination strategy.

Voting. Voting or majority voting has been used for centuries by humans to make decisions. The same methodology is employed to determine the final outcome of multiple classifier system. Three versions of majority voting exist [18]: unanimous voting in which all agree to the final decision, simple majority voting in which the final decision exceeds 50% + 1 votes, and plurality voting in which one with the most votes becomes the final decision. While these approaches cast entire vote to a single class that each classifier considers most likely, voting can also combine classifiers by averaging each classifier's probability estimates. We average each classifier's probability estimates when using voting strategy in our experiment.

Stacking. Stacking or stacked generalization [19] uses a high-level method (called level-1 generalizer or meta-learner) to combine lower-level methods (called level-0 models or base classifiers). The predictions of the lower-level methods are used as training data for high-level method. The ensemble learning proceeds in two steps: first, the predictions of level-0 models are calculated; then, the predictions are used as training data for training level-1 generalizer. The class labels of the original data are retained for level-1 learner's training data. In essence, stacking provides the meta-learner indirect feedback about the correctness of its base classifiers [20].

Grading. Grading [20] uses "graded" predictions (i.e., whether the prediction is correct, marked as "+", or incorrect, marked as "-") of the base classifiers to train a meta-classification scheme. The learning proceeds in two-steps first by obtaining correct/incorrect predictions of the base classifiers, and then replacing original class values with correct/incorrect predictions to learn meta-classification schemes. When a new instance is tested, each base classifier makes a prediction, and the final prediction is derived from the base classifiers that are predicted to be correct by the meta-classification schemes. If conflicts exist within several base-level predictions, they can be resolved by voting and other methods.

Existing researches have mainly applied ensemble-based system to credit scoring analysis [21-24], bankruptcy prediction [25], heart disease diagnosis [26, 27], and traffic incident detection [28]. Some have focused on location prediction [29]. We also apply our GBN-assisted ensemble approach to the problem of location prediction.

Table 1. Variables, variable values, and number of values in the location prediction dataset

<table>
<tr><th>Variable</th><th>Variable Value (No. of Values)</th></tr>
<tr><td>Location Departed</td><td>600thAnniversaryBuilding, BasketballCourt, Bicheondang, BusinessBuilding, CentralLibrary, DasanHallOfEconomics, EastGate, FacultyHall, FrontGate, GeumjandiSquare, HoamHall, InternationalHall, LargePlayground, LawBuilding, Myeongnyundang, Oacknyujeong, OutsideCampus, RearGate, StudentUnion, SuseonHall, SuseonHallAnnex, ToegyeHallOfHumanities, Yanghyeongwan, Yurimhoegwan (24)</td></tr>
<tr><td>Path Start</td><td>A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, Y (24)</td></tr>
<tr><td>Path Middle</td><td>A, AB, ABG, B, BA, BC, BE, BF, BFJ, BG, BGI, BGJ, BGM, BH, BHI, BQ, BQJ, C, CB, CBG, D, E, EBG, F, FB, G, GB, GBA, GJ, GM, H, HB, I, IJ, IM, J, JF, JFB, JG, JGB, JI, JK, JQB, K, M, MGB, MJ, N, NJ, none, Q, QB, QBA, SGB, T, X, XM (57)</td></tr>
<tr><td>Path End</td><td>A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y (25)</td></tr>
<tr><td>Location Arrived</td><td>600thAnniversaryBuilding, BasketballCourt, Bicheondang, BusinessBuilding, CentralLibrary, DasanHallOfEconomics, EastGate, FacultyHall, FrontGate, GeumjandiSquare, HoamHall, InternationalHall, InternationalHouse, LargePlayground, LawBuilding, Oacknyujeong, OutsideCampus, RearGate, StudentUnion, SuseonHall, SuseonHallAnnex, ToegyeHallOfHumanities, Yanghyeongwan, Yurimhoegwan (24)</td></tr>
<tr><td>Activity</td><td>ChatWithFriends, ClubActivity, Consult, Eat, Exercise, FinancialErrands, Hobby, Homework, InternetSearch, JobHunting, Lecture, MiscErrands, Other, Part-TimeJob, Shop, Study, TeaTime (17)</td></tr>
<tr><td>Major</td><td>BizAdmin, Confucianism, DomesticScience, Economics, Education, Engineering, FineArts, FreeMajor, InfoTechnology, Law, LiberalArts, SocialScience, SportsScience (13)</td></tr>
<tr><td>Student Year</td><td>Freshman, Sophomore, Junior, Senior (4)</td></tr>
<tr><td>Gender</td><td>Male, Female (2)</td></tr>
<tr><td>Weekday Leisure</td><td rowspan="2">Concert/Exhibitions, Games, IndividualSports, Socialize, TeamSports, Travel (6)</td></tr>
<tr><td>Lunch Time Leisure</td></tr>
<tr><td>Monthly Allowance</td><td><100, 100-300, 300-500, 500-700, 700-900, >900 (6)</td></tr>
</table>

3 Empirical Evaluation

To investigate the location prediction performance of the GBN-assisted ensemble system, we collected real-world data from undergraduate students, constructed two-classifier and three-classifier ensemble systems, and evaluated their performances. Hereafter, we indicate ensemble systems as ensembles or ensemble classifiers.

3.1 Data

User context data were collected from 335 college students in Seoul, Korea to create training and test data for location prediction experiment. There were 205 male students and 130 female students, and the student year composition was as follows: 131 freshmen, 38 sophomores, 64 juniors, and 102 seniors. The students were asked to complete a demographic survey which asked his/her gender, major, student year, weekday leisure activity, lunch-time leisure activity, monthly allowance, and student ID. Then, they were instructed to document their whole-day activity on campus for any two days; in particular, where they visited via what route and what activity they engaged in at the visited location. A list of campus location codes, route codes (a list of letters was specified to describe a sequence of paths), and activity codes were provided to help them record their activities. They were given extra credits for their work.

After the two types of data (i.e., the demographic data and the campus activity data) were cleaned, they were merged using the student ID to create a campus activity-demographic data. The merged data contained 12 variables: 'Location Arrived', 'Path Start', 'Path Middle', 'Path End', 'Location Departed', 'Activity', 'Gender', 'Major', 'Year', 'Weekday Leisure', 'Lunch Time Leisure', and 'Monthly Allowance'. Table 1 lists the 12 variables and the values of each variable. A total of 3,150 records of campus activity-demographic data were used in the experiment. The 'Location Arrived' variable was selected as the class variable in the experiment.

3.2 Experimental Setup

WEKA [30], an open source data-mining tool, was used to construct and evaluate the GBN-based ensembles and the GBN-assisted ensembles. The GBN-based ensembles were built using all 12 variables of the campus activity-demographic data whereas the GBN-assisted ensembles used only 5 ('Location Arrived', 'Path End', 'Location Departed', 'Activity', and 'Major'); these 5 variables were selected on the basis of the Markov blanket of GBN's class node. That is, prior to creating the GBN-assisted ensembles, a GBN was created using the 12-variable data to identify the variables that parsimoniously describe the class variable. Note that the class variable is included in the 12-variable and 5-variable dataset.

Three types of decision trees (ID3 [31], C4.5 [32] or J48 in WEKA, and CART [33]), two types of Bayesian network classifiers (GBN-K2 [34] and GBN-Hill Climb, hereafter GBN-HC), and one SVM [35] were integrated to create two-classifier (GBN+DecisionTree or GBN+SVM) and three-classifier (GBN+DecisionTree+SVM) ensemble classifiers. Each ensemble classifier employed voting, stacking, and grading strategy. All in all, 24 two-classifier (Table 3) and 18 three-classifier (Table 4) ensembles were created for each of the GBN-based and the GBN-assisted approaches.

All algorithms needed for creating the individual classifiers and the ensemble classifiers were already available in WEKA version 3.6.2. For the GBN-K2 and the GBN-HC classifier construction, the maximum number of parent node was set to 2 and the BAYES scoring metric was used. For the decision trees (ID3, J48, CART) and the SVM (SMO algorithm), the default settings in WEKA were used. As for the three combination strategies, the averaging of the probability estimates was used to combine classifiers for voting. For level-1 generalizer (or meta-classifier) for stacking, instance-based learning algorithm (IB*k* [36]) was used with ten nearest neighbors following [20]. The same meta-classification scheme was used for grading.

We performed one run of 10-fold cross-validation on each ensemble classifier to obtain its prediction accuracy, and then conducted paired t-tests at the 1% and 5% significance level to compare each ensemble classifier with the baseline GBN-K2 individual classifier.

3.3 Results

Table 2 compares the prediction accuracies of individual classifiers created using the 12 variables (first row) and 5 variables (second row). The SVM shows the best individual-classifier prediction accuracy in both cases; ID3 shows the worst. Compared to the GBN-K2 individual classifier (84.38), only the SVM created using 5 variables shows significantly better performance at the 5% significance level.

Table 2. Prediction accuracies of 12-variable and 5-variable individual classifiers compared to GBN-K2 individual classifier. (* $p<0.05$)

Classifier	GBN-K2	GBN-HC	ID3	J48	CART	SVM
12-variable	84.38	81.46	80.67	84.73	83.97	85.14
5-variable	84.38	84.79	79.14	83.08	83.71	**85.52***

Table 3 compares the prediction accuracies of the GBN-based (12 variables are used) two-classifier ensembles versus the GBN-assisted (5 variables inside the Markov blanket of the GBN's class node are selected and used) two-classifier ensembles. The prediction accuracies showing statistically better performance to the GBN-K2 individual classifier are marked in bold in the tables. Compared to the GBN-K2 individual classifier, the GBN-K2+ID3 voted-ensemble (85.56, 5-variable) and the GBN-K2+J48 graded-ensemble (85.90, 12-variable) show significantly better performance at the 1% significance level; six and three additional ensemble classifiers show better performance at the 5% significance level for the GBN-based (12-variable) and the GBN-assisted (5-variable) approaches, respectively. Overall, the GBN+ID3 voted-ensembles and GBN+SVM graded-ensembles show good performances. It is notable that GBNs can benefit from the lowest-performing ID3 individual classifier; conversely, some classifier combination can hurt the performance as shown in the case of the GBN-K2+SVM stacked-ensemble (84.22 and 83.78).

Table 3. Prediction accuracies of GBN-based (12-variable) and GBN-assisted (5-variable) 2-classifier ensembles compared to GBN-K2 individual classifier. (* p<0.05, ** p<0.01)

2-Classifier Ensemble		Voting		Stacking		Grading		Average	
		12-var	5-var	12-var	5-var	12-var	5-var	12-var	5-var
GBN-K2	ID3	**85.97***	**85.56****	85.17	84.98	84.67	84.41	85.27	84.98
	J48	**85.68***	84.48	**85.84***	84.51	**85.90****	84.98	85.81	84.66
	Cart	84.63	84.41	84.00	83.75	84.79	84.57	84.47	84.24
	SVM	84.48	84.44	84.22	83.78	**85.62***	**85.11***	84.77	84.44
GBN-HC	ID3	**85.97***	**85.62***	84.51	84.95	82.92	84.35	84.47	84.97
	J48	84.60	84.10	84.03	84.95	84.22	84.98	84.28	84.68
	Cart	84.22	84.73	84.06	84.16	84.10	84.70	84.13	84.53
	SVM	81.56	84.79	82.48	84.35	**85.43***	**85.56***	83.16	84.90
Average		84.64	84.77	84.29	84.43	84.71	84.83	84.54	84.68

Table 4. Prediction accuracies of GBN-based (12-variable) and GBN-assisted (5-variable) 3-classifier ensembles compared to GBN-K2 individual classifier. (* p<0.05, ** p<0.01)

3-Classifier Ensemble		Voting		Stacking		Grading		Average	
		12-var	5-var	12-var	5-var	12-var	5-var	12-var	5-var
GBN-K2 + SVM	ID3	**86.00***	**85.52****	85.30	84.63	**86.16****	85.17	85.82	85.11
	J48	**85.71***	84.48	85.56	84.70	**85.90****	**85.71****	85.72	84.96
	Cart	84.67	84.41	84.10	83.68	**85.87***	**85.30***	84.88	84.46
GBN-HC + SVM	ID3	**85.97***	**85.59***	85.05	84.98	**85.87***	85.24	85.63	85.27
	J48	84.60	84.16	83.94	84.51	**85.87****	**85.62****	84.80	84.76
	Cart	84.22	84.73	83.56	83.52	85.49	85.08	84.42	84.44
Average		85.20	84.82	84.59	84.34	85.86	85.35	85.21	84.84

Table 4 shows the prediction accuracies of the GBN-based (12-variable) three-classifier ensembles versus the GBN-assisted (5-variable) three-classifier ensembles. The prediction accuracy increases in general for the graded ensemble approach when an SVM classifier is added to the GBN+DecisionTree ensembles. Compared to the GBN-K2 individual classifier, three GBN-based and three GBN-assisted ensembles show significantly better prediction performance at the 1% significance level and five GBN-based and two GBN-assisted ensembles show better performance at the 5% significance level. The voting strategy seems to work well with the ID3-included ensembles and the grading strategy with the J48-included ensembles regardless of the GBN-based or GBN-assisted approaches. On the contrary, stacking, as a combination strategy, does not seem to work well; of the twelve GBN+DecisionTree+SVM stacked-ensembles, seven show worse performance than their GBN+DecisionTree two-classifier counterparts. For instance, the GBN-HC+J48+SVM and GBN-HC+CART+SVM ensembles both display lower prediction performance than their GBN+DecisionTree two-classifier counterparts regardless of the GBN-based or GBN-assisted approaches. In general, the grading strategy turned out to be the best ensemble combination strategy for the location prediction dataset in this paper, the voting strategy turned out to be the second best, and the stacking strategy turned out to be the worst strategy.

4 Discussion

In the experiment, we were able to confirm that the prediction accuracies of 24 ensemble classifiers (Tables 3 and 4, numbers in bold) were statistically better than the baseline GBN-K2 individual classifier. But how do they compare to one another? To see whether any of these superior ensemble classifiers perform significantly better than the other, we conducted a one-way ANOVA using a significance level of $\alpha = 0.05$ to check the statistical differences in prediction accuracy. The result showed that the differences among the prediction accuracies of 24 ensemble classifiers were statistically insignificant ($F(23, 216) = .209$, $p = 1.000$); no one ensemble classifier outperformed the other. With no clear winner present, choosing a two-classifier ensemble (Table 3, any one in bold) over a three-classifier ensemble (Table 4, any one in bold) could be a rational decision since the two-classifier ensembles require lower computational cost while maintaining comparable prediction performance to the three-classifier ensembles. Similar rationale could be applied to the GBN-based (12-variable) approach versus the GBN-assisted (5-variable) approach; with less number of features (variables) to work on, choosing the GBN-*assisted* approach could save on computational cost without hurting the prediction performance; but this is not the only advantage of the GBN-assisted approach.

Because the GBN-assisted ensemble approach exploits the Markov blanket of the target node, context prediction can be performed efficiently by focusing on the truly relevant explanatory variable(s); the GBN-assisted ensemble approach can be said to encapsulate a natural feature selection capability that identifies the features that parsimoniously describe the target variable. Such feature selection (or reduction) capability is beneficial to both humans and machines. For example, when designing a context prediction system, data gathering strategy for context prediction must be coordinated. If the prediction model inside the system requires too many context data to predict future context, both the users and the system will need to make much effort in handling these data. Keeping the number of features to a minimum lowers the data-handling cost for both the users and machines, and it also keeps the model simple.

Although the ensemble systems sacrifice model interpretability over performance, both the GBN-based ensemble approach and the GBN-assisted ensemble approach allow humans to understand the variable relationship through individual GBN models. A GBN expresses the relationship between a target variable and explanatory variables using nodes and links; humans can easily interpret how variables influence each other through this graph model. Since humans can understand which explanatory variables directly influence the target variable in the GBN, the graph model can be used in what-if and goal-seeking analyses. A what-if analysis is one in which decision makers analyze the possible results by considering intended changes to input conditions. A goal-seeking analysis is closely related to such simulation activities in which a certain goal is suggested, and decision makers attempt to observe what kind of input conditions are necessary to obtain such a goal. Such capabilities are the advantages of the GBN-related ensemble approaches. Lastly, it should be noted that although the stacking approach discussed in this paper showed a poor performance, changing the meta-level learner and parameter setting may improve the performance.

5 Concluding Remarks

A GBN-assisted ensemble system exploits the Markov blanket of the GBN's target node to identify and select the core features. We compared the prediction performance of the GBN-assisted approach, which uses fewer variables, to the GBN-based approach, which uses more variables, and found that the performance of the two approaches were comparable to each other despite the fact that the GBN-assisted approach handled fewer features. In this sense, we can view the GBN-assisted ensemble approach to have a computational edge over the GBN-based ensemble approach. Ensemble systems generally have better prediction accuracy, but have low interpretability of the models. The GBN-related ensemble approaches can compensate for the sacrificed interpretability by exploiting and exploring the variable relationship depicted in the individual GBN graph model. In the future, we plan to design a context prediction system which can handle multiple prediction models suited to various user groups; we plan to place the GBN-assisted ensemble systems at the heart of the system.

Acknowledgments. This research was supported by the Ubiquitous Computing and Network (UCN) Project, Knowledge and Economy Frontier R&D Program of the Ministry of Knowledge Economy (MKE) in Korea as a result of UCN's subproject 09C1-T2-20S.

References

1. Kwon, O., Yoo, K., Suh, E.: UbiDSS: A Proactive Intelligent Decision Support System as an Expert System Deploying Ubiquitous Computing Technologies. Expert Systems with Applications 28, 149–161 (2005)
2. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing 5, 4–7 (2001)
3. Patterson, D., Liao, L., Fox, D., Kautz, H.: Inferring High-Level Behavior from Low-Level Sensors. In: Dey, A.K., Schmidt, A., McCarthy, J.F. (eds.) UbiComp 2003. LNCS, vol. 2864, pp. 73–89. Springer, Heidelberg (2003)
4. Hwang, K.S., Cho, S.B.: Landmark Detection from Mobile Life Log Using a Modular Bayesian Network Model. Expert Systems with Applications 36, 12065–12076 (2009)
5. van Kasteren, T., Kröse, B.: Bayesian Activity Recognition in Residence for Elders. Intelligent Environments, 209–212 (2007)
6. Sánchez, D., Tentori, M., Favela, J.: Activity Recognition for the Smart Hospital. IEEE Intelligent Systems 23, 50–57 (2008)
7. Byun, H.E., Cheverst, K.: Utilizing Context History to Provide Dynamic Adaptations. Applied Artificial Intelligence 18, 533–548 (2004)
8. Lum, W.Y., Lau, F.C.M.: A Context-Aware Decision Engine for Content Adaptation. IEEE Pervasive Computing 1, 41–49 (2002)
9. Matsuo, Y., Okazaki, N., Izumi, K., Nakamura, Y., Nishimura, T., Hasida, K., Nakashima, H.: Inferring Long-Term User Properties Based on Users' Location History. In: 20th International Joint Conference on Artificial Intelligence (IJCAI), pp. 2159–2165 (2007)
10. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo (1988)
11. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)

12. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco (2005)
13. Kittler, J.: Combining Classifiers: A Theoretical Framework. Pattern Analysis & Applications 1, 18–27 (1998)
14. Polikar, R.: Ensemble Based Systems in Decision Making. IEEE Circuits and Systems Magazine 6, 21–45 (2006)
15. Breiman, L.: Bagging Predictors. Machine Learning 24, 123–140 (1996)
16. Schapire, R.E.: The Strength of Weak Learnability. Machine Learning 5, 197–227 (1990)
17. Ho, T.K.: The Random Subspace Method for Constructing Decision Forests. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, 832–844 (1998)
18. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms. Wiley-Interscience, New Jersey (2004)
19. Wolpert, D.H.: Stacked Generalization. Neural Networks 5, 241–259 (1992)
20. Seewald, A., Fürnkranz, J.: An Evaluation of Grading Classifiers. In: Advances in Intelligent Data Analysis, pp. 115–124 (2001)
21. Hsieh, N.C., Hung, L.P.: A Data Driven Ensemble Classifier for Credit Scoring Analysis. Expert Systems with Applications 37, 534–545 (2010)
22. Zhou, L., Lai, K.K., Yu, L.: Least Squares Support Vector Machines Ensemble Models for Credit Scoring. Expert Systems with Applications 37, 127–133 (2010)
23. Twala, B.: Multiple Classifier Application to Credit Risk Assessment. Expert Systems with Applications 37, 3326–3336 (2010)
24. Yu, L., Yue, W., Wang, S., Lai, K.K.: Support Vector Machine Based Multiagent Ensemble Learning for Credit Risk Evaluation. Expert Systems with Applications 37, 1351–1360 (2010)
25. Hung, C., Chen, J.H.: A Selective Ensemble Based on Expected Probabilities for Bankruptcy Prediction. Expert Systems with Applications 36, 5297–5303 (2009)
26. Das, R., Sengur, A.: Evaluation of Ensemble Methods for Diagnosing of Valvular Heart Disease. Expert Systems with Applications 37, 5110–5115 (2010)
27. Eom, J.H., Kim, S.C., Zhang, B.T.: AptaCDSS-E: A Classifier Ensemble-Based Clinical Decision Support System for Cardiovascular Disease Level Prediction. Expert Systems with Applications 34, 2465–2479 (2008)
28. Chen, S., Wang, W., Van Zuylen, H.: Construct Support Vector Machine Ensemble to Detect Traffic Incident. Expert Systems with Applications 36, 10976–10986 (2009)
29. Anagnostopoulos, T., Anagnostopoulos, C., Hadjiefthymiades, S., Kyriakakos, M., Kalousis, A.: Predicting the Location of Mobile Users: A Machine Learning Approach. In: Int'l Conf. Pervasive Services, pp. 65–72 (2009)
30. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA Data Mining Software: An Update. ACM SIGKDD Explorations Newsletter 11, 10–18 (2009)
31. Quinlan, J.R.: Induction of Decision Trees. Machine Learning 1, 81–106 (1986)
32. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo (1993)
33. Breiman, L.: Classification and Regression Trees. Chapman & Hall/CRC (1984)
34. Cooper, G.F., Herskovits, E.: A Bayesian Method for the Induction of Probabilistic Networks from Data. Machine Learning 9, 309–347 (1992)
35. Platt, J.C.: Fast Training of Support Vector Machines Using Sequential Minimal Optimization. In: Advances in Kernel Methods: Support Vector Learning, pp. 185–208. MIT Press, Cambridge (1999)
36. Aha, D.W., Kibler, D., Albert, M.K.: Instance-Based Learning Algorithms. Machine Learning 6, 37–66 (1991)