Rule Extraction from Support Vector Machine Using Modified Active Learning Based Approach: An Application to CRM

M.A.H. Farquad[1,2], V. Ravi[1,*], and S. Bapi Raju[2]

[1] Institute for Development and Research in banking Technology, Castle Hills Road #1, Masab Tank, Hyderabad – 500 057 (AP) India
[2] Department of Computer & Information Sciences, University of Hyderabad, Hyderabad – 500 046 (AP) India
farquadonline@gmail.com, rav_padma@yahoo.com, bapics@uohyd.ernet.in

Abstract. Despite superior generalization performance Support vector machines (SVMs) generate *black box* models. The process of converting such opaque models into transparent model is often regarded as *rule extraction*. This paper presents a new approach for rule extraction from SVMs using modified active learning based approach (mALBA), to predict churn in bank credit cards. The dataset is obtained from Business Intelligence Cup 2004, which is highly unbalanced with 93% loyal and 7% churned customers' data. Since identifying churner is paramount from business perspective, therefore considering sensitivity alone, the empirical results suggest that the proposed rule extraction approach using mALBA yielded the best sensitivity compared to other classifiers.

Keywords: Support Vector Machines, Rule Extraction, modified active learning based approach, Customer Churn.

1 Introduction

Data mining involves the use of sophisticated data analysis algorithms to discover previously unknown, valid patterns and relationships in large datasets [1-4]. Data mining algorithms consists of statistics or machine learning based approaches, such as neural networks, decision trees etc. Similarly, the validity of the patterns discovered is dependent on how they compare to "*real world*" circumstances. Data mining has been effectively applied in wide range of applications, such as fraud detection [5] and scientific discovery [6] and manufacturing [7].

In recent past it is observed that, banks and the service industries has become more customer centric. The problem of customers shifting loyalties from one organization to another is called "*churn*", and is common nowadays. Hence, there is a pressing need to develop algorithmic models that can predict which existing '*loyal*' customer is going to churn out in near future [8]. Customer Relationship Management (CRM) is

[*] Corresponding author. Ph: +91-40-2353 4981.

R. Setchi et al. (Eds.): KES 2010, Part I, LNAI 6276, pp. 461–470, 2010.

a process or methodology used to learn more about customers' needs and behaviors in order to develop stronger relationships with them. Research shows that, the customers with longer time relationship with the firm are more profitable [9, 10] than online bank customers [11]. Management should prepare an anti-churn strategy that is usually far less expensive than acquiring new customers [12, 13]. There are several types of CRM introduced for different purposes; *Operational CRM, Analytical CRM, Collaborative CRM* and *web-based CRM* [14-16].

Over a decade researchers have applied machine learning techniques for churn prediction problem; such as Multivariate Regression Analysis [17], Logistic Regression [9], Neural Networks [18], Random Forest [19], Decision Tree [20], FuzzyARTMAP [21], Support Vector Machines [22] and ensemble systems [8]. Yu Zhao et al., [23] concluded that using improved one-class SVM has shown best performance compared to other traditional methods like ANN, Decision Tree, and Naïve Bayes. SVM's generalization ability to deal with noisy data is reported using news paper subscription churn prediction data [24]. The efficiency of SVM has extended to predict the churn in Commercial Bank's VIP Customers [25]. Cao et al. reported that SVM-RFE extracts less key attributes and exhibits better satisfactory predictive effectiveness [26]. Further, recently SVMs efficiency is analyzed for customer churn prediction in landline telecommunications [27].

In this paper we present a modified active learning based approach for rule extraction from SVM using NBTree rule induction technique. The proposed approach is an extension to the approach presented by Martens et al., [41], where they used the training set with extra generated samples and using C4.5 rules were generated. During our proposed approach, support vectors are first obtained from SVM and using mALBA synthetic data instances are generated. The generated synthetic data is appended to support vectors and the target values are then replaced by the predictions of SVM. This modified data is fed to NBTree to generate rules.

This paper is structured as follows: In section 2 previous SVM rule extraction techniques are discussed. Section 3 describes the proposed approach. Next, in section 4, dataset description and experimental setup is detailed. Section 5 presents the results and discussions, and section 6 concludes this paper.

2 Rule Extraction from SVM

Gallant [28] initiated the work of rule extraction from a neural network that defines the knowledge learnt in the form of *if-then* rules. Even limited explanation can positively influence the system acceptance by the user [29]. A learning system might discover salient features in the input data whose importance was not previously recognized [30]. Rule extraction from opaque models improves generalization.

SVMs [31] have proved to be good alternative compared to other machine learning techniques specifically for classification problems [32]. Even though SVMs work well, it is completely non-intuitive to human experts, that they do not let us know the knowledge learnt by them during training in simple, comprehensible and transparent way.

Recently attempts have been made to extract rules from SVMs to represent the knowledge learnt by SVM during training. Extensive work was done towards developing rule extraction techniques for neural networks [33] but less work is reported towards rule extraction from SVM. SVM+Prototype [34], RulExtSVM [35], Extracting rules from trained support vector machines [36], Hyper rectangle Rules Extraction (HRE) [37], Fuzzy Rule Extraction (FREx) [38], Multiple Kernel-Support Vector Machine (MK-SVM) [39], SQRex-SVM [40], Active Learning-Based Approach (ALBA) [41], Hybrid rule extraction technique [42, 22, 43] and recently regression rule extraction technique [44], are some of the approaches proposed towards rule extraction from SVM.

3 Proposed Rule Extraction Approach

In this work, we propose a modified active learning based rule extraction procedure to extract rules from SVM using NBTree (Naive Bayes Tree) [45]. The proposed approach is applied to predict churn in bank credit cards. The dataset is obtained from Business Intelligence Cup 2004. The dataset is median scale and is highly unbalanced with 93% loyal and 7% churned customers' data. The proposed modified active learning based approach is described in Algorithm 1.

Algorithm 1: mALBA for Rule Extraction from SVM

Step 1: Train SVM and obtain the support vectors using training data [41].

Calculate the average distance $dist_k$ of training data to support vectors, in each dimension k

Step 2: Calculate the $dist_k$ between support vectors and training instances, in each dimension k. $dist_k = dist_k + |d_k - sv_{j,k}|$

Modified ALBA

Step 3: Randomly generate an extra data instance x_i close to support vectors

For i = 1 to 500/1000 **do**

For $k = 1$ to n **do**

$$x_{i,k} = sv(j,k) + \left[(2 * rand - 1) \times \frac{dist_k}{2} \right]$$

End for

Append the generated data to the support vectors

Step 4: Provide a class label y_i using the trained SVM as oracle.

End for

Rule generation and evaluation

Step 5: Run rule induction algorithm on the modified data set and evaluate the performance of rules in terms of accuracy, fidelity and number of rules.

SVM model is first developed using training set under 10-fold cross validation and support vectors are extracted. We have chosen an RBF kernel for developing SVM model, as it is shown to achieve good overall performance [22]. Next, synthetic data is generated using mALBA. Before generation the data instances the distance *dist* between support vectors and the training set is calculated. Using the $dist_k$, the new data instances are generated which are near support vectors. 500 and 1000 data instances are generated for empirical analysis. Generated data is then appended to the support vectors set and the predictions are obtained and the actual target values are then replaced by the predictions of SVM. This modified data is then fed to NBTree [45] to generate rules. The proposed approach is depicted in Fig. 1.

The current study in this paper is different from ALBA [41] approach in several ways, such as;

- They generated the instances using $[(rand-0.5)*dist_k/2]$, which generates the data near SVs that are between -0.5 to 0.5, whereas we are using $[(2*rand-1)\times dist_k/2]$, which generates the data near SVs that are between -1 to 1.
- They appended the generated data to training set whereas we are appending the generated data to the support vectors set.
- They employed C4.5 and RIPPER algorithm for rule generation, whereas we employed NBTree algorithm for rule generation.
- They evaluated their proposed approach for small scale and balanced problems whereas we analyzed medium scale and unbalanced problems in this study.

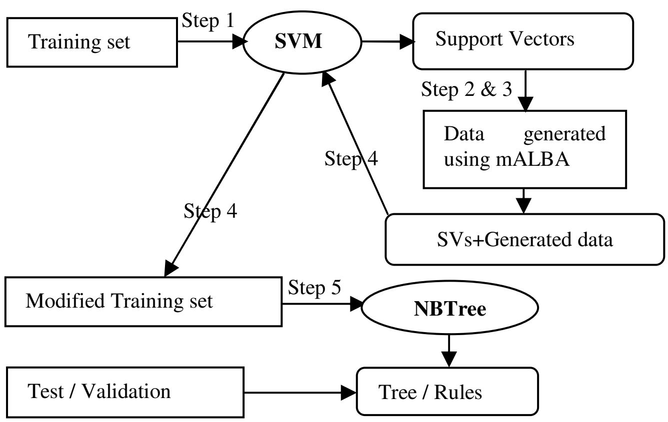

Fig. 1. Block diagram of the proposed rule extraction approach

4 Dataset Description and Experimental Setup

The dataset is from a Latin American bank that suffered from an increasing number of churns with respect to their credit card customers and decided to improve its retention system. The attribute information is tabulated in Table 1. The dataset consists of 14814 records, of which 13812 are loyal customers i.e. 93% and 1002 are churners i.e. 7%. Hence, the dataset is highly unbalanced in terms of the proportion of churners versus non-churners [46].

4.1 Experimental Setup

The available large scale unbalanced dataset is first divided into two parts of 80:20 ratios. 80% of the data is then used for training under 10 fold cross validation. 20% of the data is named as validation set and stored for evaluating the efficiency of the rules generated under 10-FCV. The efficiency and validity of the rules generated during 10-FCV are then tested against the validation set, which is a subset of the original data.

To compare the performances with the original ALBA, we have applied the rule induction techniques

1. On ALBA.
2. On ALBA with support vectors set i.e. ALBA(SVs).
3. On mALBA.

Table 1. Feature description of churn prediction data set

Feature	Description	Value
Target	Target Variable	0-NonChurner 1-Churner
CRED_T	Credit in month T	Positive real number
CRED_T-1	Credit in month T-1	Positive real number
CRED_T-2	Credit in month T-2	Positive real number
NCC_T	Number of credit cards in months T	Positive integer value
NCC_T-1	Number of credit cards in months T-1	Positive integer value
NCC_T-2	Number of credit cards in months T-2	Positive integer value
INCOME	Customer's Income	Positive real number
N_EDUC	Customer's educational level	1 - University student 2 - Medium degree 3 - Technical degree 4 - University degree
AGE	Customer's age	Positive integer
SX	Customers sex	1 - male 0 - Female
E_CIV	Civilian status	1-Single 2-Married 3-Widow 4-Divorced
T_WEB_T	Number of web transaction in months T	Positive integer
T_WEB_T-1	Number of web transaction in months T-1	Positive integer
T_WEB_T-2	Number of web transaction in months T-2	Positive integer
MAR_T	Customer's margin for the company in months T	Real Number
MAR_T-1	Customer's margin for the company in months T-1	Real Number
MAR_T-2	Customer's margin for the company in months T-2	Real Number
MAR_T-3	Customer's margin for the company in months T-3	Real Number
MAR_T-4	Customer's margin for the company in months T-4	Real Number
MAR_T-5	Customer's margin for the company in months T-5	Real Number
MAR_T-6	Customer's margin for the company in months T-6	Real Number

5 Results and Discussions

Identifying potential churners correctly is the basic intension of many business decision makers. Hence, they place high emphasis on sensitivity alone which contributes towards the bottom-line of the fundamental CRM. Consequently in this paper, sensitivity is accorded top priority ahead of specificity and accuracy. We used the SVM library viz., LibSVM [47] for SVM. LibSVM is integrated software for support vector classification and is developed in MATLAB. RapidMiner4.5 community edition [48] is used for generating NBTree. The quantities employed to measure the quality of the classifiers are sensitivity, specificity and accuracy [49].

During our empirical study we generated 500 and 1000 extra instances separately, using the calculations described in step 3 of the proposed approach and using original ALBA calculations. The results obtained using NBTree and Decision Tree are presented in Table 2 and 3, respectively.

Table 2. Average Results obtained using NBTree

Extra Generated Data	Classifiers	10-Fold Cross validation				Validation			
		Sens*	Spec*	Acc*	t-test	Sens*	Spec*	Acc*	t-test
	SVM	63.35	81.41	80.19	4.93	64.65	80.63	79.55	4.79
500	**ALBA**	65.48	85.23	83.92	3.44	67.7	84.52	83.38	3.13
	ALBA (SVs)	74.93	83.05	82.5	0.82	76.55	82.74	82.32	0.76
	mALBA	**78.17**	**80.36**	**80.28**	-	**79.35**	**79.16**	**79.17**	-
1000	**ALBA**	65.35	85.88	84.46	2.2	68.05	84.75	83.68	2.45
	ALBA (SVs)	73.8	83.26	82.62	0.11	73.25	82.99	82.27	0.69
	mALBA	**74.3**	**82.84**	**82.31**	-	**75.9**	**83.3**	**82.87**	-

Table 3. Average Results obtained using Decision Tree

Extra Generated Data	Classifiers	10-Fold Cross validation				Validation			
		Sens*	Spec*	Acc*	t-test	Sens*	Spec*	Acc*	t-test
500	**ALBA**	60.22	83.56	81.97	4.377	65.2	82.65	81.48	3.647
	ALBA (SVs)	69.68	80.38	79.66	1.089	72.7	80.38	72.25	0.875
	mALBA	**73.81**	**76.91**	**76.7**	-	**75.05**	**75.99**	**76.43**	-
1000	**ALBA**	61.6	83.81	82.32	2.874	64.05	83	81.72	3.211
	ALBA (SVs)	67.69	80.55	79.68	0.658	71.65	79.7	79.17	0.418
	mALBA	**70.1**	**81.6**	**80.85**	-	**72.85**	**81.3**	**80.7**	-

It is observed that the rules extracted by the proposed rule extraction approach using mALBA with 500 extra instances yielded best average sensitivity of 78.17% under 10-FCV and the same set of rules yielded 79.35% sensitivity against validation set. Using ALBA, the sensitivity yielded under 10-FCV is 65.48% and against validation set the sensitivity obtained is 67.7%. It is observed from the results that the generated instances using mALBA are positively near the SVM boundary. And the extra generated data using ALBA when used with SVs set, the sensitivity yielded is 74.93% during 10-FCV and against validation it yielded 76.55% sensitivty.

The same set of experiments is carried out by generating 1000 extra instances, it is observed that mALBA yielded 74.3% sensitivity, whereas original ALBA yielded 65.35% sensitivity under 10-FCV. mALBA and ALBA with 1000 extra instances yielded 75.9% and 68.05% sensitivities against validation set, respectively. It is observed that the time taken by ALBA for rule extraction is more than mALBA, as the extra generated instances are appended to the training set in ALBA approach and the extra generated instances are appended to SVs set in mALBA. When the data is generated using ALBA calculations with SVs the sensitivity yielded under 10-FCV is 73.8% and against validation set the sensitivity yielded is 73.25%. It is observed that the generated samples with SVs yielded better sensitivity compared to ALBA. When the generated data is used with SVs, the complexity, time and rules are decreased. It is observed that instead of using all the training instances with the generated data for rule induction algorithm as [41], it is better to take SVs set with extra generated data to reduce the complexity of the system and it also produces less number of rules without compromising the accuracy of the model.

For comparison purpose rules are also extracted using DT. It is observed that rules extracted using NBTree yielded best sensitivity compared to the rules extracted using DT. Furthermore, the average number of rules during 10-FCV using NBTree is 13 whereas the average number of rules extracted using DT is 85.5 for mALBA, 400 for ALBA and 98.5 for ALBA (SVs).

Using sensitivity, the classifiers are compared with t-test at $n_1+n_2-2=10+10-2=18$ degrees of freedom at 10% level of significance. We tested if the difference in performances is statistically significant. The tabulated value of t-statistics for 18 degrees of freedom at 10% level of significance is 1.73. That means, if t-statistics value between two different classifiers is more than 1.73, we say that the difference between techniques is statistically significant otherwise not significant. The t-test values obtained between the sensitivities shows that mALBA is statistically significant to original ALBA but it is statistically insignificant to ALBA with SVs.

A rule set is considered to display a high level of *fidelity* if it can *mimic* the behavior of the machine learning technique from which it was extracted i.e. SVM in our study. The fidelity obtained using ALBA, ALBA with SVs and mALBA is presented in Table 4. It is observed that ALBA behaves 83.28% like SVM with 500 generated samples, whereas our proposed mALBA approach behaves 82.65% like SVM. The fidelity obtained using 1000 extra generated instances with ALBA, mALBA and ALBA with SVs is 81.64%, 79.1% and 79.03%, respectively. It is observed that ALBA mimics the behavior of SVM better than mALBA and ALBA with SVs.

Table 4. Average Fidelity Obtained using ALBA and Proposed mALBA

Extra Generated Data	Classifiers	NBTree	DecisionTree
500	**ALBA**	83.28	89.85
	ALBA (SVs)	80.88	86.56
	mALBA	82.65	85.64
1000	**ALBA**	81.64	90.54
	ALBA (SVs)	79.03	87.78
	mALBA	79.1	86.37

Table 5 presents the example rule set obtained using mALBA. The number of rules extracted using mALBA is very much less in number when compared to the rules extracted using ALBA [41]. It is observed that mALBA with 500 extra samples yielded the best sensitivity among other approaches tested and more generalized rules are obtained.

Table 5. Rules Obtained using mALBA

S. No	Antecedents	Consequent
1	If MAR_T <= 8.81 and MAR_T-5 <= -5.04	Non-Churner
2	If MAR_T <= 8.81 and MAR_T-5 > -5.04	Churner
3	If MAR_T > 8.81 and CRED_T < 594.145	Churner
4	If MAR_T > 8.81 and CRED_T > 594.145 and CRED_T-2 < 94.58	Churner
5	If MAR_T > 8.81 and CRED_T > 594.145 and CRED_T-2 > 94.58	Non-Churner
6	If CRED_T<=598.1 and MAR-T_2<=14.045 and MAR-t_5<=973	Churner

6 Conclusions

In this paper, we present a modified active learning based approach for rule extraction from SVM to solve credit card customer churn prediction problem. The dataset is taken from Business Intelligence Cup organized by University of Chile in 2004. This is highly unbalanced data with 93% good customers and 7% churned customers. While solving the problems like churn prediction, sensitivity is accorded high priority. Accordingly, by considering sensitivity alone, it is observed that the proposed rule extraction approach using mALBA yielded the best sensitivity of 79.35%. It is also observed that when ALBA is used with SVs set obtained better accuracy than original ALBA [41]. It is observed that original ALBA using C4.5 for generating rules, yielded more number of rules, which may make the black box model a transparent model but the comprehensibility of the classifier is adversely affected. Efficiency of the feature selection using SVM-RFE can be analyzed in future works. Further, mALBA rule extraction approach can be employed to solve Insurance fraud detection problem.

References

1. Usama, F., Piatetsky-Shapiro, G., Smyth, P.: From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence (1996)
2. Adriaans, P., Zantinge, D.: Data Mining. Addison-Wesley, New York (1996)
3. Osuna, E.E., Freund, R., Girosi, F.: Support vector Machines: Training and Applications. Technical Report, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, AI Memo No. 1602 (1997)
4. Osuna, E.E., Freund, R., Girosi, F.: Training support vector machines: an application to face detection. In: Proceedings of Computer Vision and Pattern Recognition, pp. 130–136 (1997)
5. Senator, T., Goldberg, H.G., Wooton, J., Cottini, M.A., Umarkhan, A.F., Klinger, C.D., Llamas, W.M., Marrone, M.P., Wong, R.W.H.: The Financial Crimes Enforcement Network AI System (FAIS): Identifying Potential Money Laundering from Reports of Large Cash Transactions. AI Magazine 16(4), 21–39 (1995)
6. Hall, J., Mani, G., Barr, D.: Applying Computational Intelligence to the Investment Process. In: Proceedings of CIFER 1996: Computational Intelligence in Financial Engineering, Washington, D.C. IEEE Computer Society Press, Los Alamitos (1996)
7. Ravi, V., Arul Shalom, S.A., Manickavel, A.: Sputter Process Variables Prediction via Data Mining. In: Proceedings of the 2004 IEEE Conference on Cybernetics and Intelligent Systems, Singapore (2004)
8. Kumar, D.A., Ravi, V.: Predicting credit card customer churn in banks using data mining. International Journal for Data Analysis, Techniques and Strategies 1(1), 4–28 (2008)
9. Bolton, R.N., Kannan, P.K., Bramlett, M.D.: Implications of loyalty program membership and service experiences for customer retention and value. Journal of the Academy of Marketing Science 28(1), 95–108 (2000)
10. Lejeune, M.A.P.M.: Measuring the impact of data mining on churn management. Electronic Networking Applications and Policy 11(5), 375–387 (2001)
11. Mols, N.P.: The Behavioral consequences of PC banking. International Journal of Bank Marketing 16(5), 195–201 (1998)
12. Chu, B.H., Tsai, M.-S., Ho, C.-S.: Toward a hybrid data mining model for customer retention. Knowledge-Based Systems 20(8), 703–718 (2007)

13. Ryals, L., Knox, S.: Cross-functional issues in the implementation of relationship marketing through customer relationship management. European Management Journal 19(5), 534–542 (2001)
14. Alajoutsijarvi, K., Mannermaaa, K., Tikkanen, H.: Customer relationships and the small software firm: a framework for understanding challenges faced in marketing. Information and Management 37, 153–159 (2000)
15. Laudon, K., Laudon, J.: Management Information Systems: Managing the Digital Firm, 7th edn. Prentice-Hall, Englewood Cliffs (2000)
16. Edwards, J.: Get It Together with Collaborative CRM. insideCRM. Tippit (2007), `http://www.insidecrm.com/features/collaborative-crm-112907/`
17. Bloemer, J., Ruyter, K.D., Peeters, P.: Investigating drivers of bank loyalty: the complex relationship between image, service quality and satisfaction. International Journal of Bank Marketing 16(7), 276–286 (1998)
18. Mozer, M.C., Wolniewicz, R., Grimes, D.B., Johnson, E., Kaushansky, H.: Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. IEEE Transactions of Neural Networks 11(3), 690–696 (2000)
19. Larivière, B., Van den Poel, D.: Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: the case of financial services. Expert Systems with Applications 27(2), 277–285 (2004)
20. Hung, S.-Y., Yen, D.C., Wang, H.: Applying data mining to telecom churn management. Expert Systems with Applications 31(3), 515–524 (2006)
21. Naveen, N., Ravi, V., Kumar, D.A.: Application of fuzzyARTMAP for churn prediction in bank credit cards. International Journal of Information and Decision Sciences 1(4), 428–444 (2009)
22. Farquad, M.A.H., Ravi, V., Bapi, R.S.: Data Mining using Rules Extracted from SVM: an Application to Churn Prediction in Bank Credit Cards. In: Sakai, H., Chakraborty, M.K., Hassanien, A.E., Ślęzak, D., Zhu, W. (eds.) RSFDGrC 2009. LNCS, vol. 5908, pp. 390–397. Springer, Heidelberg (2009)
23. Zhao, Y., Li, B., Li, X., Liu, W., Ren, S.: Customer Churn Prediction Using Improved One-Class Support Vector Machine. In: Li, X., Wang, S., Dong, Z.Y. (eds.) ADMA 2005. LNCS (LNAI), vol. 3584, pp. 300–306. Springer, Heidelberg (2005)
24. Coussement., K., Van den Poel., D.: Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. Expert Systems with Applications 34(1), 313–327 (2008)
25. Zhaoa., J., Dang, X.-H.: Bank Customer Churn Prediction Based on Support Vector Machine: Taking a Commercial Bank's VIP Customer Churn as the Example. In: Proceedings of the 4th International Conference on Wireless Communications, Networking and Mobile Computing 2008 (WiCOM 2008), pp. 1–4 (2008)
26. Cao., K., Shao, P.-j.: Customer Churn Prediction Based on SVM-RFE. In: International Seminar on Business and Information Management ISBIM 2008, vol. 1, pp. 306–309 (2008)
27. Huang, B.Q., Kechadi, T.-M., Buckley, B., Kiernan, G., Keogh., E., Rashid., T.: A new feature set with new window techniques for customer churn prediction in land-line telecommunications. Expert Systems with Applications 37(5), 3657–3665 (2010)
28. Gallant, S.: Connectionist expert systems. Communications of the ACM 31(2), 152–169 (1988)
29. Davis, R., Buchanan, B.G., Shortliffe, E.: Production rules as a representation for a knowledge-based consultation program. Artificial Intelligence 8(1), 15–45 (1977)
30. Craven, M.W., Shavlik, J.W.: Using sampling and queries to extract rules from trained neural networks. In: The Proceedings of the Eleventh International Conference on Machine Learning, San Francisco, CA, USA (1994)
31. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer, New York (1995)

32. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.N.: Gene selection for cancer classification using support vector machines. Machine Learning 46(1-3), 389–422 (2002)
33. Tickle, A.B., Andrews, R., Golea, M., Diederich, J.: The truth will come to light: directions and challenges in extracting the knowledge embedded within trained artificial neural network. IEEE Transactions on Neural Networks 9(6), 1057–1068 (1998)
34. Nunez, H., Angulo, C., Catata, A.: Rule extraction from support vector machines. In: European Symposium on Artificial Neural Networks Proceedings, pp. 107–112 (2002)
35. Fung, G., Sandilya, S., Bharat, R.R.: Rule extraction from linear support vector machines. In: Proceeding of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, pp. 32–40. ACM Press, New York (2005)
36. Barakath, N., Diederich, J.: Eclectic rule-extraction from support vector machines. International journal of Computer Intelligence 2(1), 59–62 (2005)
37. Zhang, Y., Su, H., Jia, T., Chu, J.: Rule Extraction from Trained Support Vector Machines. In: Ho, T.-B., Cheung, D., Liu, H. (eds.) PAKDD 2005. LNCS (LNAI), vol. 3518, pp. 61–70. Springer, Heidelberg (2005)
38. Chaves, A.C.F., Vellasco, M.M.B.R., Tanscheit, R.: Fuzzy rule extraction from support vector machines. In: Fifth International Conference on Hybrid Intelligent Systems, Rio de Janeiro, Brazil, November 06-09 (2005)
39. Chen, Z., Li, J., Wei, L.: A multiple kernel support vector machine scheme for feature selection and rule extraction from gene expression data of cancer tissue. Artificial Intelligence in Medicine 41, 161–175 (2007)
40. Barakat, N.H., Bradley, A.P.: Rule Extraction from Support Vector Machines: A Sequential Covering Approach. IEEE Transactions on Knowledge and Data Engineering 19(6), 729–741 (2007)
41. Martens, D., Baesens, B., Gestel, T.V.: Decompositional Rule Extraction from Support Vector Machines by Active Learning. IEEE Transactions on Knowledge and Data Engineering 21(2), 178–191 (2009)
42. Farquad, M.A.H., Ravi, V., Bapi, R.S.: Rule Extraction using Support Vector Machine Based Hybrid Classifier. In: Presented in TENCON-2008, IEEE region 10 Conference, Hyderabad, India, November 19-21 (2008)
43. Farquad, M.A.H., Ravi, V., Bapi, R.S.: Support Vector Machine based Hybrid Classifiers and Rule Extraction Thereof: Application to Bankruptcy Prediction in Banks. In: Soria, E., Martín, J.D., Magdalena, R., Martínez, M., Serrano, A.J. (eds.) Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods and Techniques, vol. 2, pp. 404–426. IGI Global, USA (2010)
44. Farquad, M.A.H., Ravi, V., Bapi, R.S.: Support vector regression based hybrid rule extraction methods for forecasting. Expert Systems with Applications (2010), doi:10.1016/j.eswa.2010.02.055
45. Ron, K.: Scaling Up the Accuracy of Naïve-Bayes Classifiers: a Decisoin-Tree Hybrid. In: Proceedings of KDD 1996, Portland, USA (1996)
46. Business Intelligence Cup - 2004: Organized by the University of Chile, `http://www.tis.cl/bicup_04/text-bicup/BICUP/202004/20public/20data.zip`
47. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines (2001), `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
48. Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., Euler, T.: YALE: Rapid Prototyping for Complex Data Mining Tasks. In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2006 (2006)
49. Fawcett, T.: An introduction to ROC analysis. Pattern Recognition Letters 27, 861–874 (2006)