Journal of Theoretical and Applied Information Technology
15th October 2016. Vol.92. No.1
© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ENSEMBLE SELECTION AND OPTIMIZATION BASED ON SOFT SET THEORY FOR CUSTOMER CHURN CLASSIFICATION

***[1]MOHD KHALID AWANG, [1]MOKHAIRI MAKHTAR, [1]M NORDIN A RAHMAN, [2]MUSTAFA MAT DERIS**

[1]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin,

22000 Tembila, Terengganu, Malaysia

[2] Faculty of Computer Science and Information Technology,

Universiti Tun Hussein Onn Malaysia (UTHM)

Batu Pahat, Johor, Malaysia

*Corresponding author: khalid@unisza.edu.my

ABSTRACT

Ensemble methods or multiple classifiers which combine decisions from many base classifiers have been confirmed to outperform the classification performance of any single classifiers. Despite having the ability of producing the highest classification accuracy, ensemble methods have suffered significantly from their large volume of base classifiers. Thus, in the previous work, we have proposed a novel soft set based method to prune the base classifiers from heterogeneous ensemble committee and have demonstrated the ability of our proposed soft set pruning algorithm in reducing a substantial number of classifiers while at the same time producing the highest prediction accuracy. However, the pruning method only suggests a subset of relevant classifiers, and the search for the optimized and best classifiers is not yet considered. The selection of the best or optimized classifiers is carried out by checking all combinations of pruned classifiers. In this paper, we extended our research by proposing a new soft ensemble selection and optimization method to find the best subset of the pruned classifiers. The results of this work have proven that our proposed method is able to search for the minimum number of classifiers in the ensemble repository while at the same time maintaining or improving the classification performance. The proposed method is systematically evaluated using Customer Churn dataset taken from the UC Irvine Machine Learning Repository data set. This work proved that the proposed soft ensemble selection and optimization method is able to search for the minimum number of classifiers in the ensemble repository while at the same time improving the classification performance.

Keywords: *Ensemble Selection, Customer Churn Prediction, Ensemble Optimization, Soft Set, Ensemble Methods*

1. INTRODUCTION

Customer churn prediction is one of the most significant activities that form the foundation for all Customer Relationship Management (CRM) in any business operations. Maintaining current customers is vital since the cost of acquiring new customers is relatively high [28]. Retain current customers is important because the cost of getting new customers is very high [28]. In the telecommunications industry, customer churn become more apparent because of the fast growth of wireless technology. Customers have more choices and able to select and change from one package to another package presented by the different service providers. This action is known as churn and now has turned into a major concern of network providers [29]. Therefore, to remain competitive in an increasingly saturated market, companies will strive to retain their customers and try to reduce the cost of acquiring new customers. Many researchers have tried to establish various models of customer churn prediction, for example the neural network, decision tree, genetic algorithm, and regression analysis [30-32]. However, current approaches for churn prediction

Journal of Theoretical and Applied Information Technology
15th October 2016. Vol.92. No.1
© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

are not effective and need to be improved because there are many uncertainty factors that contribute to customer churn.

Ensemble methods or multiple classifiers are known as a new learning algorithm that train a set of base classifiers and combine their results to accomplish the best prediction accuracy [1]. Previous researches have revealed that combining the predictions of a collection of classifiers can be an effective strategy to improve classification performance, such as bagging [2], boosting [3], stacking [4], Bayes optimal classifier [5], rotation forest [5], ensemble selection [6] and hybrid intelligent system [7].

Generally, ensemble methods consist of two main stages: the production of multiple base classifier models and their combination [36]. The nature of ensemble methods which is to produce a large number of individual classifiers tends to have drawbacks. One of the biggest threats of ensemble method is the resource consumption.

Our previous work [40] proposed a novel approach for an ensemble pruning method based on the soft set theory. This new approach aims to solve the problem of representing less redundant ensemble classifiers based on the dimensionality reduction of soft set theory. Thus, in the previous work, we have proposed a novel soft set based method to prune the classifiers from heterogeneous ensemble committee and demonstrated the ability of our proposed soft set ensemble pruning to reduce a substantial number of classifiers and at the same time producing the highest prediction accuracy. However, the pruning method only suggests a subset of relevant classifiers, and the search for the optimized and best classifiers is not yet considered. The selection of the best or optimized classifiers is carried out by checking all combinations of pruned classifiers. Thus, this paper extends our previous work by proposing a new method for ensemble classifier selection and optimization from the pruned ensemble.

The rest of this paper is organized as follows. Section 2 describes the ensemble methods and ensemble selection. Section 3 discusses the soft set and its reduction algorithm. Section 4, soft set ensemble selection and optimization method. Section 5 describes the experimental setting and results. Finally, Section 6 summarizes this work.

2. ENSEMBLE SELECTION AND OPTIMIZATION

Previous researchers have proposed various ensemble methods as learning algorithms in data mining to improve the classifiers performance and accuracy. There is no single ensemble methods that dominate classification technique. Most of the previous studies focus on the ensemble construction and ensemble combination in improving the accuracy and performance of classification, but rarely consider the ensemble pruning and optimization algorithms. Nevertheless, there are few researches focusing on ensemble pruning and optimization methods [8-13, 36, 40-41]. In general, ensemble methods could be explained as in Figure 1.

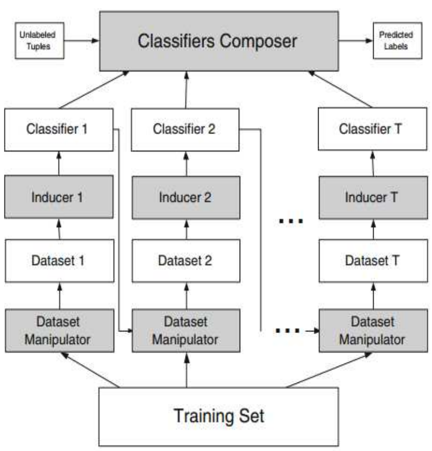

Figure 1: Ensemble Framework (Rokach 2010)

In Figure 1, [37] describes the typical ensemble framework method for classification tasks. The framework is divided into 4 different phases which are:

1. Training set
2. Base inducer
3. Diversity Generator
4. Combiner or Composer

The ensemble started with the selection of the dataset for the training set. After the training set is selected, the next phase is generating the base inducer or ensemble construction where the classification algorithms are selected and trained with the training data set. The diversity generator will ensure the diversity of the base classifiers. Finally the diverse classifiers are combined to form the final ensemble.

© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In our previous work [40], we have proposed soft set pruning ensemble framework. We have introduced the soft ensemble pruning to eliminate the similar or redundant classifiers in ensemble pool. In this paper, we further enhance the framework by implementing the soft ensemble selection and optimization method based on soft set reduction algorithms. The final stage of our framework is the combination of classifiers. Based on the previous researches, there are numerous combination approaches, such as majority voting, weighted majority voting, summation, product, maximum and minimum, fuzzy integral, Dempster-Shafer based fusion or decision templates [38, 39]. In this research, we employed the simple majority voting as our combination approach.

3. SOFT SET THEORY

Soft set is a parametrized general mathematical tool which deals with a collection of approximate descriptions of objects. Soft set theory has potential applications in many different fields which include the smoothness of functions, game theory, operations research, Riemann integration, Perron integration, probability theory, and measurement theory, attribute and feature reduction [22-25]

a) Basic Concept of Soft Set

Let U be initial universal set and let E be a set of parameters. Let $P(U)$ denote the power set of U. A pair (F, E) is called a soft set over U, if only if F is a mapping given by $F{:}E \rightarrow P(U)$ [22,23].

b) Soft Set Based Reduction based on Discernibility

The most fundamental concept in rough set is set approximation and it is carried out by discernibility matrices and discernibility functions. Based on [24] that every rough set is a soft set, we proposed a similar concept of discernibility function in rough set [25,26] to reduct and discern the soft set data. In this paper, we have developed our soft ensemble selection and optimization algorithm based on dimensionality reduction by A. Skrowron [25].

4. PROPOSED SOFT ENSEMBLE SELECTION AND OPTIMIZATION METHOD

In our previous work, a new approach to reduce the classifiers in the ensemble pool has been developed. The proposed method is known as soft set ensemble pruning algorithm. The testing data set and the classifier prediction transforms into a soft set tabulation format, where the column representing the classifiers while the row representing the instances of the testing data set. Experimental results show that our proposed soft set ensemble pruning able to reduce the size of an ensemble considerably and at the same time maintaining the classification performance.

Once of the drawback in our previous work is that the whole pruned has to be searched in order to find the best or optimized ensemble classifiers. We defined an optimized classed as having the least number of classifiers in an ensemble while producing the best prediction accuracy. The process of finding the optimized ensemble starting with selecting the best classifiers from the original pruned ensemble classifiers. Then, apply the soft set reduction algorithm to the remaining of the pruned classifiers. The base classifier is combined with the new best classifiers from the new soft set reduction algorithm if the combination improves the prediction accuracy. We employ the simple majority voting technique for the combination's procedure.

A New Soft Set Ensemble Selection and Optimization Algorithm

Input: A Pruned Subset of Classifiers
Output: Optimal Set of Classifiers in Ensemble Team

1. Start
2. Transform the decision table with the Prediction and Actual Output into softest representation and become NEW REDUCT
3. Select the BEST Classifier which having the best performance (ACC) and set it as a BASE classifier
4. Generate a discernibility matrix based on the NEW REDUCT
5. Transform the discernibility matrix into discernibility function
6. Apply the absorption law on NEW REDUCT to get the set of the reduct.
7. Apply the distributive law to construct the reduct and become a NEW REDUCT
8. FOR ALL possibilities of NEW REDUCT
   a. Find the BEST Classifier from the NEW REDUCT, which having the best

Journal of Theoretical and Applied Information Technology
15th October 2016. Vol.92. No.1
© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

performance (ACC) and set it as a COMPLEMENT

b. Combine the BASE with the COMPLEMENT to become a NEW BASE

c. Apply the Simple Majority Voting technique to get the ACC of the NEW BASE

d. IF the ACC is Higher than BASE, continue step 4 to 8.

9. End

This work not only focuses on the prediction accuracy, but also the optimum size of the ensemble classifier. We proposed soft set reduction algorithm to find the optimized ensemble classifier based on the discernibility function.

5. EXPERIMENTAL EVALUATION

In order to validate the performance of the proposed soft set ensemble selection and optimization algorithm, we construct our ensemble on customer churn datasets.

Data set

The dataset is based on the customer churn taken from the UCI Repository of Machine Learning Databases. The dataset consists of 3,333 cleaned objects and 20 instances along with one indicator whether or not to churn [27]. The dataset is divided into training and testing dataset with the proportion of 2978 and 355 respectively. The dataset consists of the features as in Table 1.

Table 1 : Original Features of Customer Churn Prediction

Input Features	Data Type	Description
X_1=State	Categorical	Represent the 50 states and the district of Columbia
X_2=Account length	Numeric	The variable for how long account has been active
X_3=Area code	Categorical	Represent the area code
X_4=Phone number	Text	Essentially a surrogate key for customer identification
X_5=International Plan	Categorical	Dichotomous categorical having yes or no value
X_6=Voice Mail Plan	Categorical	Dichotomous categorical variable yes or no value
X_7=Number of voice mail messages	Numeric	Integer valued variable
X_8=Total day minutes	Numeric	Continuous variable for number of minute customer has used the service during the day
X_9=Total day calls	Numeric	Integer-valued variable
X_{10}=Total day charge	Numeric	Continuous variable based on foregoing two variables
X_{11}=Total evening minutes	Numeric	Continuous variable for number of minute customer has used the service during the evening
X_{12}=Total evening calls	Numeric	Integer-valued variable
X_{13}=Total evening charge	Numeric	Continuous variable based on previous two variables
X_{14}=Total night minutes	Numeric	Continuous variable for storing minutes the customer has used the service during the night
X_{15}=Total night calls	Numeric	Integer-valued variable
X_{16}=Total night charge	Numeric	Continuous variable based on foregoing two variables
X_{17}=Total international minutes	Numeric	Continuous variable to minute customer has used service to make international calls
X_{18}=Total international calls	Numeric	Integer-valued variable
X_{19}=Total international charge	Numeric	Continuous variable based on foregoing two variables
X_{20}=Number of calls to customer service	Numeric	Integer-valued variable
Y_1=actual result	Categorical	OUTPUT-Churn: Yes or No

Learning Algorithm for Classifiers

We create our heterogeneous ensemble by selecting ten different classifiers which are as in Table 2.

Table 2 : Classifiers as provided in Weka

Journal of Theoretical and Applied Information Technology
15th October 2016. Vol.92. No.1
© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Classifiers	Team Representation	Prediction Accuracy
weka.classifiers.meta.EnsembleSelection	0000000001	0.84
weka.classifiers.rules.DecisionTable	0000000010	0.85
weka.classifiers.meta.StackingC	0000000100	0.85
weka.classifiers.meta.AdaBoostM1	0000001000	0.85
weka.classifiers.meta.Bagging	0000010000	0.77
weka.classifiers.rules.ZeroR	0000100000	0.85
weka.classifiers.bayes.NaiveBayesUpdateable	0001000000	0.82
weka.classifiers.rules.JRip	0010000000	0.85
weka.classifiers.trees.J48	0100000000	0.85
weka.classifiers.lazy.IBk	1000000000	0.85

Table 2 displays the prediction accuracy of each of the classifiers with the highest accuracy of an individual classifier is 0.85. Based on the number of classifiers in the ensemble, we could end up with 1024 combination of different classifiers teams for the dataset as illustrated in Table 3.

Table 3 : Number of Classifier Before and After Pruning

	Original Ensemble Size	**Pruned Ensemble Size**
Number of classifiers in ensemble	10	5
Number of all possible combinations of classifiers (2^n)	2^{10} =1024	2^5=32
Percentage of reduction		96.9%

Table 3 shows the size of ensemble before and after the soft set pruning algorithm. The original set of ensemble consist 10 classifiers which is: `{c1,c2,c3,c4,c5,c6,c7,c8,c9,c10}`.

The actual size of the ensemble and the number of all possible combination of classifiers are significantly reduced from 1024 teams in 32, which is a 96.9 % reduction.

Meanwhile the soft set pruning algorithms for dataset drops 5 classifiers which are `{c2, c5, c7, c8, c9}` and produce the a new subset which is:
`{c1,c3,c4,c6,c10}`.

In the previous work, we have proposed a novel soft set based method to prune the classifiers from heterogeneous ensemble committee and demonstrated the ability of our proposed soft set ensemble pruning to reduce a substantial number of classifiers and at the same time producing the highest prediction accuracy. However, the pruning method only suggests a subset of relevant classifiers, and the search for the optimized and best classifiers is not yet considered. The selection of the best or optimized classifiers is carried out by checking all combinations of pruned classifiers. Thus, this paper extends our previous work by proposing a new method for ensemble classifier selection and optimization from the pruned ensemble.

Soft Set Ensemble Selection and Optimization algorithm

1. Start
2. BASE = {}
3. New Reduct = pruned ensemble classifier = **{c1,c3,c4,c6,c10}**
   3.1. Choose **c1** (highest ACC) as BEST
   3.2. BASE = BASE + BEST = {c1}
4. New Reduct = **{c3,c4,c6,c10}**
   4.1. Generate Discernibility Matrik on New Reduct
   4.2. Generate Discernibility Function on New Discenibilty Matrik
   4.3. Generate New Reduct
5. New Reduct = **{c3,c4,c6}**
   5.1. Choose **c3** (highest ACC) as BEST
   5.2. BASE = BASE + BEST = {c1,c3}
   5.3. New Reduct = **{c4,c6}**
   5.4. Generate Discernibility Matrik on New Reduct
   5.5. Generate Discernibility Function on New Discenibilty Matrik
   5.6. Generate New Reduct
   5.7. New Reduct = **{c4}**
6. Choose **c4** (highest ACC) and single value as Best
7. **BASE = BASE + BEST = {c1,c3,c4}**
8. End

Table 4 : Ensemble Combination with the Highest Accuracy

Best Team of	FULL ensemble	Prediction Accuracy	Number of Classifiers

© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Ensemble Classifiers	or Pruned		in the Team Ensemble
1111110000	Full [1024 teams]	0.86	6
1111101000	Full [1024 teams]	0.86	6
1111010000	Full [1024 teams]	0.86	5
1111000000	Full [1024 teams]	0.86	4
1011000000	**Pruned** [32 teams]	**0.86**	**3**

Table 4 shows some of the possible combinations of classifiers in the ensemble methods that produce the best prediction accuracy of 0.86. Based on the experiment, we could conclude that the performance of the ensemble classifiers is better than any other ensemble methods or single classifiers. Furthermore, the soft set ensemble selection and optimization algorithm able to produce the minimum number of classifiers team. The experimental result shows that the performance of the proposed soft set ensemble selection and optimization is as good as the full ensemble.

The best ensemble team constitutes of three (3) classifiers; which is 0011000001 = {c3,c4,c10}.

6. CONCLUSION

In this paper, a new soft set based ensemble selection and optimization method is proposed. Heterogeneous ensemble is generated based on ten different classifier algorithms. It's acknowledged that the most significant advantage of soft set theory is its great ability of dimensionality reduction. Based on this soft set reduction algorithm, our previous work has demonstrated that the ensemble is pruned and only a subset of the classifiers is considered prior to ensemble combination. However, the pruning method only suggests a subset of relevant classifiers, and the search for the optimized and best classifiers is not yet considered. The selection of the best or optimized classifiers is carried out by checking all combinations of pruned classifiers. In this paper, we extended our research by proposing a new soft ensemble selection and optimization method to find the optimized classifier ensemble from the pruned ensemble classifier. The results of this work have proven that our proposed method is able to search for the minimum number of classifiers in the ensemble while at the same time maintaining or improving the classification performance. The proposed method is systematically evaluated using Customer Churn dataset taken from the UCI repository. Soft Set ensemble selection and optimization not only reduce the number of members of the ensemble, but able to produce highest prediction accuracy.

ACKNOWLEDGEMENT :

This work is partially supported by UniSZA and KPT (Grant No. FRGS/2/2013/ICT07/UniSZA/02/2).

REFRENCES:

[1] Dietterich, T. G. ,"Ensemble methods in machine learning. In Multiple classifier systems, *Springer Berlin Heidelberg,* 2000*,* pp. 1-15

[2] Breiman, L. , "Bagging predictors. Machine learning," 24(2), 1996, pp. 123-140

[3] Freund, Y., & Schapire, R. E., Experiments with a new boosting algorithm," *In ICML*, 1996, July, (Vol. 96, pp. 148-156)

[4] Breiman, L., "Stacked regressions. Machine learning", 24(1), 1996, pp. 49-64

[5] Wang, H., Fan, W., Yu, P. S., & Han, J. , " Mining concept-drifting data streams using ensemble classifiers." *In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, August,2003, pp.226-235

[6] Rodriguez, J. J., Kuncheva, L. I., & Alonso, C. J. , "Rotation forest: A new classifier ensemble method. Pattern Analysis and Machine Intelligence," *IEEE Transaction*s on, 28(10), 2006, pp.1619-1630.

[7] Caruana, R., Niculescu-Mizil, A., Crew, G., & Ksikes, A "Ensemble selection from libraries of models," *In Proceedings of the twenty-first international conference on Machine learning* ACM, July 2004, pp 18

[8] Tsoumakas, G., Partalas, I., & Vlahavas, I. , " A taxonomy and short review of ensemble selection," *In Workshop on Supervised and*

© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Unsupervised Ensemble Methods and Their Applications, July 2008.

[9] Partalas, I., Tsoumakas, G., Katakis, I., & Vlahavas, I. "Ensemble pruning using reinforcement learning," *In Advances in Artificial Intelligence* Springer Berlin Heidelberg. 2006, pp.301-310

[10] Martinez-Muoz, G., Hernández-Lobato, D., & Suarez, A. , "An analysis of ensemble pruning techniques based on ordered aggregation." *Pattern Analysis and Machine Intelligence, IEEE Transactions* on, 31(2),2009, pp. 245-259.

[11] Caruana, R., Munson, A., & Niculescu-Mizil, A. , "Getting the most out of ensemble selection. In Data Mining," ICDM'06. Sixth International Conference on ,2006, pp. 828-833. IEEE.

[12] Cruz, R. M., Sabourin, R., Cavalcanti, G. D., & Ren, T. I. , META-DES: A dynamic ensemble selection framework using meta-learning. Pattern Recognition," 48(5),2015, pp. 1925-1935

[13] Taghavi, Z. S., & Sajedi, H. , " Ensemble pruning based on oblivious Chained Tabu Searches," *International Journal of Hybrid Intelligent Systems*, 12(3), 2016, pp.131-143

[14] Fürnkranz, J., & Widmer, G. , " Incremental reduced error pruning, *In Proceedings of the 11th International Conference on Machine Learning (ML-94)* , 1994, pp. 70-77

[15] Margineantu, D. D., & Dietterich, T. G. , "Pruning adaptive boosting," *In ICML*, July 1997,Vol. 97, pp. 211-218).

[16] Schapire, R. E., & Singer, Y., "BoosTexter: A boosting-based system for text categorization. Machine learning, 39(2),200, pp. 135-168.

[17] Strehl, A., & Ghosh, J. , "Cluster ensembles---a knowledge reuse framework for combining multiple partitions," *The Journal of Machine Learning Research, 3,* 2003, pp. 583-617.

[18] Topchy, A., Jain, A. K., & Punch, W. , "Clustering ensembles: Models of consensus and weak partitions. Pattern Analysis and Machine Intelligence," *IEEE Transactions* on, 27(12),2005, pp. 1866-1881.

[19] Bakker, B., & Heskes, T. , " Clustering ensembles of neural network models. Neural networks," 16(2),2005, pp 261-269.

[20] Zhang, Y., Burer, S., & Street, W. N. , " Ensemble pruning via semi-definite programming," *The Journal of Machine Learning Research*, 7,2005, pp. 1315-1338.

[21] Chen, H., Tino, P., & Yao, X. , "A probabilistic ensemble pruning algorithm" *In Data Mining Workshops, 2006. ICDM Workshops* 2006. *Sixth IEEE International Conference* IEEE,pp.878-882

[22] Molodtsov, D. , " Soft set theory—first results. Computers & Mathematics with Applications," 37(4),1999, pp. 19-31.

[23] Maji, P. K., Biswas, R., & Roy, A. " Soft set theory. Computers & Mathematics with Applications," 45(4),2003,pp. 555-562.

[24] Herawan, T., & Deris, M. M. , " A direct proof of every rough set is a soft set, " *In | 2009 Third Asia International Conference on Modelling & Simulation*, May 2009,(pp. 119-124,. IEEE

[25] Skowron, A., & Rauszer, C. , "The discernibility matrices and functions in information systems," *In Intelligent Decision Support* ,1992, pp. 331-362. Springer Netherlands

[26] Kong, Z., Gao, L., Wang, L., & Li, S. , "The normal parameter reduction of soft sets and its algorithm. Computers & Mathematics with Applications", 56(12), 2008, 3029-3037

[27] 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(1),* 2009, pp. 10-18.

[28] Christopher, M., & Peck, H. 2012. *Marketing logistics.* Routledge.

[29] Verhoef, P. C., & Lemon, K. N. 2013. Successful customer value management: Key lessons and emerging trends. *European Management Journal*, 31(1), 1-15.

[30] Ismail, M. R., Awang, M. K., Rahman, M. N. A., & Makhtar, M. 2015. A Multi-Layer Perceptron Approach for Customer Churn Prediction. *International Journal of Multimedia and Ubiquitous Engineering,* 10(7), 213-222.

[31] Awang, M. K., Rahman, M. N. A., & Ismail, M. R. 2012. Data Mining for Churn Prediction: Multiple Regression Approach. *In Computer Applications for Database, Education, and Ubiquitous Computing* (pp. 318-324). Springer Berlin Heidelberg.

[32] Coussement, K., Benoit, D. F., & Van den Poel, D. 2010. Improved marketing decision making in a customer churn prediction context

© 2005 - 2016 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

using generalized additive models. *Expert Systems with Applications*, 37(3), 2132-2143.

[33] De Bock, K. W., & Van den Poel, D. 2011. An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction. *Expert Systems with Applications*, 38(10), 12293-12301.

[34] Shaaban, E., Helmy, Y., Khedr, A., & Nasr, M. 2012. A proposed churn prediction model. *IJERA*, 2, 693-697.

[35] Gower, John Clifford. "Properties of Euclidean and non-Euclidean distance matrices." *Linear Algebra and its Applications* 67 (1985): 81-97.

[36] Makhtar, M., Awang, M. K., Rahman, M. N. A., Fadzli, S. A., & Mohamad, M. (2015). Optimizing Sensitivity And Specificity Of Ensemble Classifiers For Diabetic Patients. *Journal of Theoretical and Applied Information Technology*, *82*(2).

[37] Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*,*33*(1-2), 1-39.

[38] Kuncheva, L. I. (2004). *Combining pattern classifiers: methods and algorithms*. John Wiley & Sons.

[39] Tulyakov, S., Jaeger, S., Govindaraju, V., & Doermann, D. (2008). Review of classifier combination methods. In *Machine Learning in Document Analysis and Recognition* (pp. 361-386). Springer Berlin Heidelberg.

[40] Awang, M. K., Makhtar, M., Rahman, M. N. A., & Deris, M. (2016). A New Soft Set Based Pruning Algorithm For Ensemble Method. *Journal Of Theoretical And Applied Information Technology*, *88*(3).

[41] Makhtar, M., Yang, L., Neagu, D., & Ridley, M. (2012, March). Optimisation of classifier ensemble for predictive toxicology applications. In Computer Modelling and Simulation (UKSim), 2012 UKSim 14th International Conference on (pp. 236-241). IEEE.