Classification Techniques in Machine Learning: Applications and Issues

Aized Amin Soofi* and Arshad Awan

Department of Computer Science, Allama Iqbal Open University, Islamabad, Pakistan

Abstract: Classification is a data mining (machine learning) technique used to predict group membership for data instances. There are several classification techniques that can be used for classification purpose. In this paper, we present the basic classification techniques. Later we discuss some major types of classification method including Bayesian networks, decision tree induction, k-nearest neighbor classifier and Support Vector Machines (SVM) with their strengths, weaknesses, potential applications and issues with their available solution. The goal of this study is to provide a comprehensive review of different classification techniques in machine learning. This work will be helpful for both academia and new comers in the field of machine learning to further strengthen the basis of classification methods.

Keywords: Machine learning, classification, classification review, classification applications, classification algorithms, classification issues.

1. INTRODUCTION

Machine Learning (ML) is a vast interdisciplinary field which builds upon concepts from computer science, statistics, cognitive science, engineering, optimization theory and many other disciplines of mathematics and science [1]. There are numerous applications for machine learning but data mining is most significant among all [2]. Machine learning can mainly classified into two broad categories include supervised machine learning and unsupervised machine learning.

Unsupervised machine learning used to draw conclusions from datasets consisting of input data without labeled responses [3] or we can say in unsupervised learning desired output is not given. Supervised machine learning techniques attempt to find out the relationship between input attributes (independent variables) and a target attribute (dependent variable) [4]. Supervised techniques can further classified into two main categories; classification and regression. In regression output variable takes continuous values while in classification output variable takes class labels [5].

Classification is a data mining (machine learning) approach that used to forecast group membership for data instances [6]. Although there are variety of available techniques for machine learning but classification is most widely used technique [7]. Classification is an admired task in machine learning especially in future plan and knowledge discovery. Classification is categorized as one of the supreme studied problems by researchers of the machine learning and data mining fields [8]. A general model supervised learning (classification techniques) is shown in Figure **1**.

Although classification is well known technique in machine learning but it suffers with issues like handling missing data. Missing values in data set can cause problem during both training and classification phases. Some of potential reasons of missing data are presented in [9] includes; Non entry of record due to misconception, data recognized irrelevant at the time of entry, data removal because of deviation with other documented data and equipment malfunction.

Missing data problem can overcome by approaches [10] like; Data miners can overlook the omitting data, swap whole omitting values with an individual global constant, swap an omitting value with its feature mean for the given class, manually observe samples with omitting values and insert a feasible or probable value. In this work we will focus only on some selected classification methods.

This paper organized as following; in section 2 methodology of review is presented. Section 3 is divided into four subsections in which selected classification techniques has been discussed. In section 3.1 Logic based technique (decision tree) has been discussed. In section 3.2, statistical learning techniques (Bayesian networks) are discussed. K-Nearest neighbor classifiers are presented in section 3.3. Support Vector Machines has been discussed in section 3.4.

*Address correspondence to this author at the Department of Computer Science, Allama Iqbal Open University, Islamabad, Pakistan;
Tel: +923217680092; E-mail: aizedamin@yahoo.com

ISSN: 1814-8085 / E-ISSN: 1927-5129/17

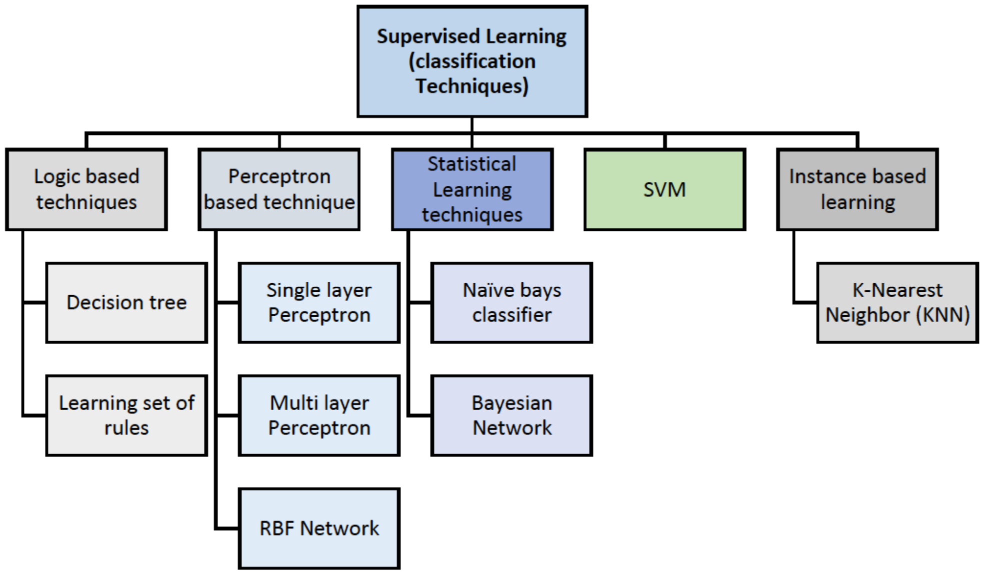

Figure 1: Supervised learning classification techniques.

2. METHODOLOGY

A literature search was performed for the articles by using databases include IEEE xplore, google scholar, science direct and some related web pages that are written in English. The keywords used for literature search include; Machine learning, data mining, classification, classification review, classification applications and classification algorithms. These keywords were used alone and in combination for the initial collection of research material. Only those articles that contain relevant data about classification techniques applications, challenges and solutions were included in this review. It is difficult to provide exhaustive review of all supervised machine learning classification methods in a single article, Therefore we focused only on commonly used classification techniques include Decision Tree (ID3 and C4.5), Bayesian Network, K-Nearest Neighbor and Support Vector Machines. Applications of different classification techniques are presented in Table **1** and issues of classification techniques with their available solutions are presented in Table **2**.

3. CLASSIFICATION TECHNIQUES

Major classification techniques has been discussed in this section with their basic working, advantages and disadvantages.

3.1. Decision Tree Induction

Decision tree algorithms are most commonly used algorithms in classification [11]. Decision tree provides an easily understandable modeling technique and it also simplifies the classification process [12]. The decision tree is transparent mechanism it facilitate users to follow a tree structure easily in order to see how the decision is made [13]. In this section basic philosophy of decision tree methods has been discussed with their strengths, limitations and applications.

The core objective of decision tree is to produce a model that calculates the value of a required variable based on numerous input variables [6]. Usually all decision tree algorithms are constructed in two phases (i) tree growth; in which training set based on local optimal criteria is splitting recursively until most of the record belonging to the partition having same class label [14] (ii) tree pruning; in which size of tree is reduced making it easier to understand [15]. In this section we will focus on ID3 and C4.5 decision tree algorithm.

ID3 (Iterative Dichotomiser 3) decision tree algorithm was introduced in 1986 [16, 17]. It is one of the widely used algorithms in the area of data mining and machine learning due to its effectiveness and simplicity [16]. The ID3 algorithm is based on information gain. Some of the strengths and weaknesses of ID3 decision tree are presented in [18], strengths includes; easy to understand and in final decision whole training example is considered while weaknesses includes; no back tracking search, unable to handle missing values and no global optimization.

C4.5 is a famous algorithm for decision trees production. It is an expansion of the ID3 algorithm and

Table 1: Classification Techniques Applications

Classification Techniques	Applications	Reference
ID3	predicting student performance	[20]
	land capability classification	[31]
	tolerance related knowledge acquisition	[32]
	computer crime forensics	[33]
	fraud detection application	[34]
C4.5	Decision making of loan application by debtor	[35]
	Predicting Software Defects	[36]
	Thrombosis collagen diseases	[37]
	Electricity price prediction	[38]
	coal logistics customer analysis	[39]
	Selecting Question Pools	[40]
Bayesian Network	automatic and interactive mode for Image Segmentation	[41]
	traffic incident detection	[42]
	signature verification	[43]
	efficient patrolling of nurses	[44]
	examine dental pain	[45]
	telecommunication and internet networks	[46]
K- Nearest neighbor	Microarray data classification	[47]
	Phoneme Prediction	[48]
	Face recognition	[49]
	Agarwood oil quality grading	[50]
	Classification of nuclear receptors and their subfamilies	[51]
	Short-term traffic flow forecasting	[52]
	Plant Leaf Recognition	[53]
SVM	Scene classification	[54]
	Predict corporate financial distress	[55]
	Induction motors fault diagnosis	[56]
	Analog circuit fault diagnosis	[57]
	enterprise market competition	[58]

it minimize its drawbacks caused by ID3. In pruning phase C4.5 tries to eliminate the un-comfort branches by swapping them with leaf nodes by going back through the tree once it has been generated [19]. The strengths of C4.5 are dealing training data with missing feature values, deals both discrete and continuous features and providing facility of both pre and post pruning [18, 20]. The weaknesses includes; not suitable for small data set [18] and high processing time as compare to other decision trees.

3.2. Bayesian Networks

A Bayesian Network (BN) refers graphical model for probability associations betwixt a set of variables [21]. BN structure S consist directed acyclic graph (DAG) and the nodes in S are in one-to-one communication with the X features. The arcs exemplify unexpected impacts betwixt the nodes while the scarcity of possible arcs in S encodes conditional liberties [2]. Normally Bayesian Network learning tasks can be isolated into two subtasks; (a) network DAG structure learning, (b) parameters determination.

One of the problems with Bayesian networks classifier is that it usually requires continuous attributes to be discretized. The process of conversion of continuous attribute into discrete attribute introduced classification issues [22, 23]. These issues may include noise, missing information and consciousness to the change of the attributes towards class variables [24]. The other method of Bayesian network classifier in which continuous attribute does not converted into discrete attribute, needs valuation of the attribute's conditional density [23].

To overcome the problem of conditional density estimation of attributes, in [24] Gaussian kernel function with stable constraints for evaluation of attributes density was used. Then Experiment was

Table 2: Classification Techniques Issue and Solutions

Classification Approach	Issue	Solution/technique	Ref.
Decision tree (ID3 and C4.5)	multi valued attributes Complex information entropy and attribute with more values Noisy data classification	Algorithm by combining ID3 and association function(AF) modification to the attribute selection methods, pre pruning strategy and rainforest approach Enhanced algorithm with Taylor formula Credal-C4.5 tree	[62] [63] [64] [65]
Bayesian Network	Attributes conditional density estimation Inference (large domain discrete and continuous variables) Multi-dimensional data	Gaussian kernel function decision-tree structured conditional probability greedy learning algorithm	[24] [66] [67]
K nearest neighbor	space requirement time requirement KNN scaling over multimedia dataset	Prototype selection feature selection and extraction methods finding R-Tree index multimedia KNN query processing system	[68] [69] [70] [30]
SVM	controlling the false positive rate low sparse SVM classifier multi-label classification	Risk Area SVM (RA-SVM) Cluster Support Vector Machine (CLSVM) fuzzy SVMs (FSVMs)	[71] [72] [73]

performed on data set given at UCI machine learning repository indicate that continuous attributes provides better classification accuracy as compare to other techniques by using Gaussian kernel function in Bayesian Network classifiers.

Some of the advantages of Bayesian network are presented in [25] includes (i) smoothness properties; minor changes in Bayesian network model do not influence the working of the system (ii) Flexible applicability; identical Bayesian Network model can be used for resolving both regression and classification issues (iii) handling missing data; Bayesian network has capability to filled out missing data by assimilating over all opportunities of the missing values.

3.3. K- Nearest Neighbor

In K-nearest neighbor (KNN) technique, nearest neighbor is measured with respect to value of k, that define how many nearest neighbors need to be examine to describe class of a sample data point [26]. Nearest neighbor technique is divided into two categories i.e, structure based KNN and structure less KNN. The structure based technique deals with the basic structure of the data where the structure has less mechanism which associated with training data samples [27]. In structure less technique entire data is categorized into sample data point and training data, distance is calculated between sample points and all training points and the point with smallest distance is known as nearest neighbor [28].

One of the main advantage of KNN technique is that it is effective for large training data and robust to noisy training data [29]. Scaling KNN queries over enormous high dimensional multimedia datasets is a stimulating issue for KNN classifiers. To overcome this issue an high performance multimedia KNN query processing system [30] was introduced, in this system the fast distance based pruning methods are coupled with suggested Distance-Pre computation based R-tree (DPR-Tree) index structure. Input/output cost is reduced by this exclusive coupling but it increase the computational work of KNN search.

Two important obstacles with nearest neighbor based classifiers are highlighted in [59] that includes; space requirement and its classification time. Different methods have been introduced to overcome space requirement issue. K-Nearest Neighbor Mean Classifier (k-NNMC) was introduced in [59]. K-NNMC independently search k nearest neighbors for every training pattern class and calculate mean for all given k-neighbors. It is presented experimentally by using numerous standard data-sets that the classification accuracy of suggested classifier is better as compare to other classifiers like weighted k-nearest neighbor classifier (Wk-NNC) [60] and it has ability to combine efficiently with any space reduction and indexing methods.

The advantages of KNN include simplicity, transparency, Robust to noisy training data, easy to understand and implement and disadvantages includes computation complexity, memory limitation, poor run-time performance for large training set and irrelevant attributes can cause problems [28, 61].

3.4. Support Vector Machines

Vapnik proposed statistical learning theory based machine learning method which is known as Support vector machine (SVM) [74]. SVM has considered as one of the highest prominent and convenient technique for solving problems related to classification of data [75] and learning and prediction[76]. Support vectors are the data points that lie closest to the decision surface [77]. It executes the classification of data vectors by a hyper plane in immense dimensional space [78]. Maximal margin classifier is the simplest or basic form of SVM that helps to determine the most simple classification problem of linear separable training data with binary classification [27]. The maximal margin classifier used to find the hyper plane with maximal margin in real world complications [79].

The main advantage of SVM is its capability to deal with wide variety of classification problems includes high dimensional and not linearly separable problems. One of the major drawback of SVM that it requires number of key parameters to set correctly to attain excellent classification results [80].

4. CONCLUSION

In this paper various popular classification techniques of machine learning has been discussed with their basic working mechanism, strengths and weaknesses. The potential applications and issues with their available solutions have also been highlighted. Classification methods are typically strong in modeling interactions. The discussed classification techniques can be implemented on different type of data set i.e. health, financial etc. It is difficult to find out which technique is superior to other because each technique has its own merits, demerits and implementation issues. The selection of classification technique depends on user problem domain. However, lot of work has been done in classification domain but it still requires formal attention of research community to overcome classification issues that have been arising due to dealing with new classification problems like problems in classification of Big Data.

REFERENCES

[1] Ghahramani Z. "Unsupervised learning," in Advanced lectures on machine learning, ed: Springer, 2004; pp. 72-112. https://doi.org/10.1007/978-3-540-28650-9_5

[2] Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. ed, 2007.

[3] Zhang D, Nunamaker JF. Powering e-learning in the new millennium: an overview of e-learning and enabling technology. Information Systems Frontiers 2003; 5: 207-218. https://doi.org/10.1023/A:1022609809036

[4] Maimon O, Rokach L. Introduction to supervised methods, in Data Mining and Knowledge Discovery Handbook, ed: Springer, 2005 pp. 149-164.

[5] Ng A. "CS229 Lecture notes."

[6] Kesavaraj G, Sukumaran S. A study on classification techniques in data mining. in Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on, 2013; pp. 1-7.

[7] Singh M, Sharma S, Kaur A. Performance Analysis of Decision Trees. International Journal of Computer Applications 2013; 71.

[8] Baradwaj BK, Pal S. Mining educational data to analyze students' performance. arXiv preprint arXiv:1201.3417, 2012.

[9] Dunham MH. Data mining: Introductory and advanced topics: Pearson Education India, 2006.

[10] Kantardzic M. Data mining: concepts, models, methods, and algorithms: John Wiley & Sons, 2011.

[11] Twa MD, Parthasarathy S, Roberts C, Mahmoud AM, Raasch TW, Bullimore MA. Automated decision tree classification of corneal shape. Optometry and vision science: official publication of the American Academy of Optometry 2005; 82: 1038. https://doi.org/10.1097/01.opx.0000192350.01045.6f

[12] Brodley CE, Utgoff PE. Multivariate versus univariate decision trees: Citeseer, 1992.

[13] Jang J-SR. ANFIS: adaptive-network-based fuzzy inference system. Systems, Man and Cybernetics, IEEE Transactions on, 1993; 23: 665-685. https://doi.org/10.1109/21.256541

[14] Rutkowski L, Pietruczuk L, Duda P, Jaworski M. Decision trees for mining data streams based on the McDiarmid's bound. Knowledge and Data Engineering, IEEE Transactions on, 2013; 25: 1272-1279. https://doi.org/10.1109/TKDE.2012.66

[15] Patil DD, Wadhai V, Gokhale J. Evaluation of decision tree pruning algorithms for complexity and classification accuracy, 2010.

[16] Quinlan JR. Induction of decision trees. Machine learning 1986; 1: 81-106. https://doi.org/10.1007/BF00116251

[17] Quinlan JR. Simplifying decision trees. International Journal of man-Machine Studies 1987; 27: 221-234. https://doi.org/10.1016/S0020-7373(87)80053-6

[18] Sharma S, Agrawal J, Agarwal S. Machine learning techniques for data mining: A survey, in Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference on, 2013; pp. 1-6.

[19] Bhukya DP, Ramachandram S. Decision tree induction: an approach for data classification using AVL-tree. International Journal of Computer and Electrical Engineering 2010; 2: 660. https://doi.org/10.7763/IJCEE.2010.V2.208

[20] Adhatrao K, Gaykar A, Dhawan A, Jha R, Honrao V. Predicting Students' Performance using ID3 and C4. 5 Classification Algorithms, arXiv preprint arXiv:1310.2071, 2013.

[21] Phyu TN. Survey of classification techniques in data mining, in Proceedings of the International MultiConference of Engineers and Computer Scientists 2009; pp. 18-20.

[22] Yang Y, Webb GI. Discretization for naive-Bayes learning: managing discretization bias and variance. Machine learning 2009; 74: 39-74. https://doi.org/10.1007/s10994-008-5083-5

[23] Friedman N, Goldszmidt M. Discretizing continuous attributes while learning Bayesian networks, in Icml 1996; pp. 157-165.

[24] Wang S-C, Gao R, Wang L-M. Bayesian network classifiers based on Gaussian kernel density. Expert Systems with Applications, 2016.

[25] Myllymäki P. Advantages of Bayesian Networks in Data Mining and Knowledge Discovery Available: http://www.bayesit.com/docs/advantages.html

[26] Cover T, Hart P. Nearest neighbor pattern classification. IEEE Transactions on Information Theory 1967; 13: 21-27. https://doi.org/10.1109/TIT.1967.1053964

[27] Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, *et al.* Top 10 algorithms in data mining. Knowledge and Information Systems 2008; 14: 1-37. https://doi.org/10.1007/s10115-007-0114-2

[28] Bhatia N. Survey of nearest neighbor techniques. arXiv preprint arXiv:1007.0085, 2010.

[29] Teknomo K. Strengths and weaknesse of K Nearest Neighbor. Available: http://people.revoledu.com/kardi/tutorial/KNN/Strength%20and%20Weakness.htm

[30] Li H, Liu L, Zhang X, Wang S. Hike: A High Performance kNN Query Processing System for Multimedia Data, in 2015 IEEE Conference on Collaboration and Internet Computing (CIC), 2015; pp. 296-303. https://doi.org/10.1109/CIC.2015.44

[31] Kumar N, Obi Reddy G, Chatterjee S, Sarkar D. An application of ID3 decision tree algorithm for land capability classification. Agropedology 2013; 22: 35-42.

[32] Shao X, Zhang G, Li P, Chen Y. Application of ID3 algorithm in knowledge acquisition for tolerance design. Journal of Materials Processing Technology 2001; 117: 66-74. https://doi.org/10.1016/S0924-0136(01)01016-0

[33] Tan Y, Qi Z, Wang J. Applications of ID3 algorithms in computer crime forensics, in Multimedia Technology (ICMT), 2011 International Conference on, 2011; pp. 4854-4857.

[34] Zou K, Sun W, Yu H, Liu F. ID3 Decision Tree in Fraud Detection Application, in Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on, 2012; pp. 399-402.

[35] Amin RK, Indwiarti, Sibaroni Y. Implementation of decision tree using C4.5 algorithm in decision making of loan application by debtor (Case study: Bank pasar of Yogyakarta Special Region), in Information and Communication Technology (ICoICT), 2015 3rd International Conference on, 2015; pp. 75-80.

[36] Li B, Shen B, Wang J, Chen Y, Zhang T. A Scenario-Based Approach to Predicting Software Defects Using Compressed C4.5 Model, in Computer Software and Applications Conference (COMPSAC), 2014 IEEE 38th Annual, 2014; pp. 406-415.

[37] Soliman SA, Abbas S, Salem ABM. Classification of thrombosis collagen diseases based on C4.5 algorithm, in 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS), 2015; pp. 131-136. https://doi.org/10.1109/IntelCIS.2015.7397209

[38] Hehui Q, Zhiwei Q. Feature selection using C4.5 algorithm for electricity price prediction, in 2014 International Conference on Machine Learning and Cybernetics, 2014; pp. 175-180. https://doi.org/10.1109/ICMLC.2014.7009113

[39] Duan F, Zhao Z, Zeng X. Application of Decision Tree Based on C4.5 in Analysis of Coal Logistics Customer, in Intelligent Information Technology Application, 2009. IITA 2009. Third International Symposium on, 2009; pp. 380-383.

[40] Seet AM, Zualkernan IA. An Adaptive Method for Selecting Question Pools Using C4.5, in 2010 10th IEEE International Conference on Advanced Learning Technologies, 2010; pp. 86-88.

[41] Zhang L, Ji Q. A Bayesian network model for automatic and interactive image segmentation, Image Processing, IEEE Transactions on, 2011; 20: 2582-2593. https://doi.org/10.1016/j.trc.2006.11.001

[42] Zhang K, Taylor MA. Effective arterial road incident detection: a Bayesian network based algorithm. Transportation Research Part C: Emerging Technologies 2006; 14: 403-417.

[43] Xiao X, Leedham G. Signature verification using a modified Bayesian network. Pattern Recognition 2002; 35: 983-995. https://doi.org/10.1016/S0031-3203(01)00088-7

[44] Aoki S, Shiba M, Majima Y, Maekawa Y. Nurse call data analysis using Bayesian network modeling, in Aware Computing (ISAC), 2010 2nd International Symposium on, 2010; pp. 272-277.

[45] Chattopadhyay S, Davis RM, Menezes DD, Singh G, Acharya RU, Tamura T. Application of Bayesian classifier for the diagnosis of dental pain, Journal of Medical Systems 36: 2012; 1425-1439. https://doi.org/10.1007/s10916-010-9604-y

[46] Bashar A, Parr G, McClean S, Scotney B, Nauck D. Knowledge discovery using Bayesian network framework for intelligent telecommunication network management, in Knowledge Science, Engineering and Management, ed: Springer, 2010; pp. 518-529.

[47] Kumar M, Rath SK. Microarray data classification using Fuzzy K-Nearest Neighbor, in Contemporary Computing and Informatics (IC3I), 2014 International Conference on, 2014; pp. 1032-1038.

[48] Rizwan M, Anderson DV. Using k-Nearest Neighbor and Speaker Ranking for Phoneme Prediction, in Machine Learning and Applications (ICMLA), 2014 13th International Conference on, 2014; pp. 383-387.

[49] Kasemsumran P, Auephanwiriyakul S, Theera-Umpon N. Face recognition using string grammar fuzzy K-nearest neighbor, in 2016 8th International Conference on Knowledge and Smart Technology (KST), 2016; pp. 55-59.

[50] Ismail N, Rahiman MHF, Taib MN, Ali NAM, Jamil M, Tajuddin SN. The grading of agarwood oil quality using k-Nearest Neighbor (k-NN), in Systems, Process & Control (ICSPC), 2013 IEEE Conference on, 2013; pp. 1-5.

[51] Tiwari AK, Srivastava R. Feature based classification of nuclear receptors and their subfamilies using fuzzy K-nearest neighbor, in Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances in, 2015; pp. 24-28.

[52] Li S, Shen Z, Xiong G. A k-nearest neighbor locally weighted regression method for short-term traffic flow forecasting, in 2012 15th International IEEE Conference on Intelligent Transportation Systems, 2012; pp. 1596-1601.

[53] Munisami T, Ramsurn M, Kishnah S, Pudaruth S. Plant Leaf Recognition Using Shape Features and Colour Histogram with K-nearest Neighbour Classifiers. Procedia Computer Science 2015; 58: 740-747. https://doi.org/10.1016/j.procs.2015.08.095

[54] Mandhala VN, Sujatha V, Devi BR. Scene classification using support vector machines, in Advanced Communication Control and Computing Technologies (ICACCCT), 2014 International Conference on, 2014; pp. 1807-1810.

[55] Zhao Y, Zhu S, Yu J, Wang L. Predicting corporate financial distress by PCA-based support vector machines, in 2010 International Conference on Networking and Information Technology, 2010; pp. 373-376. https://doi.org/10.1109/ICNIT.2010.5508491

[56] Aydin I, Karakose M, Akin E. Artificial immune based support vector machine algorithm for fault diagnosis of induction motors, in Electrical Machines and Power Electronics, 2007. ACEMP '07. International Aegean Conference on, 2007; pp. 217-221.

[57] Yehui L, Yuye Y, Liang H. Fault diagnosis of analog circuit based on support vector machines, in Communications Technology and Applications, 2009. ICCTA '09. IEEE International Conference on, 2009; pp. 40-43.

[58] Jialong H, Yanbin W. Classification of the enterprise market competition based on support vector machines, in 2010 Chinese Control and Decision Conference, 2010; pp. 1644-1647. https://doi.org/10.1109/CCDC.2010.5498321

[59] Viswanath P, Sarma TH. An improvement to k-nearest neighbor classifier, in Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE, 2011; pp. 227-231. https://doi.org/10.1109/RAICS.2011.6069307

[60] Dudani SA. The Distance-Weighted k-Nearest-Neighbor Rule. IEEE Transactions on Systems, Man, and Cybernetics, 1976; SMC-6: 325-327. https://doi.org/10.1109/TSMC.1976.5408784

[61] Cunningham P, Delany SJ. k-Nearest neighbour classifiers, 2007.

[62] Chen J, Luo D-l, Mu F-X. An improved ID3 decision tree algorithm, in Computer Science & Education, 2009. ICCSE '09. 4th International Conference on, 2009; pp. 127-130.

[63] Thakur D, Markandaiah N, Raj DS. Re optimization of ID3 and C4.5 decision tree, in Computer and Communication Technology (ICCCT), 2010 International Conference on, 2010; pp. 448-450.

[64] Huang M, Niu W, Liang X. An improved Decision Tree classification algorithm based on ID3 and the application in score analysis, in 2009 Chinese Control and Decision Conference, 2009; pp. 1876-1879. https://doi.org/10.1109/CCDC.2009.5192865

[65] Mantas CJ, Abellán J. Credal-C4.5: Decision tree based on imprecise probabilities to classify noisy data, Expert Systems with Applications 2014; 41: 4625-4637. https://doi.org/10.1016/j.eswa.2014.01.017

[66] Mori J, Mahalec V. Inference in hybrid Bayesian networks with large discrete and continuous domains. Expert Systems with Applications 2016; 49: 1-19. https://doi.org/10.1016/j.eswa.2015.11.019

[67] Hobæk Haff I, Aas K, Frigessi A, Lacal V. Structure learning in Bayesian Networks using regular vines. Computational Statistics & Data Analysis 2016; 101: 186-208. https://doi.org/10.1016/j.csda.2016.03.003

[68] Babu VS, Viswanath P. Rough-fuzzy weighted k-nearest leader classifier for large data sets. Pattern Recognition 2009; 42: 1719-1731. https://doi.org/10.1016/j.patcog.2008.11.021

[69] Duda RO, Hart PE, Stork DG. Pattern classification: John Wiley & Sons, 2012.

[70] Guttman A. R-trees: a dynamic index structure for spatial searching 1984; 14: ACM.

[71] Moraes D, Wainer J, Rocha A. Low false positive learning with support vector machines. Journal of Visual Communication and Image Representation 2016; 38: 340-350. https://doi.org/10.1016/j.jvcir.2016.03.007

[72] Carrizosa E, Nogales-Gómez A, Romero Morales D. Clustering categories in support vector machines, Omega.

[73] Abe S. Fuzzy support vector machines for multilabel classification. Pattern Recognition 2015; 48: 2110-2117. https://doi.org/10.1016/j.patcog.2015.01.009

[74] Vapnik VN. The Nature of Statistical Learning Theory, 1995.

[75] Nizar A, Dong Z, Wang Y. Power utility nontechnical loss analysis with extreme learning machine method. Power Systems, IEEE Transactions on, 2008; 23: 946-955. https://doi.org/10.1109/TPWRS.2008.926431

[76] Xiao H, Peng F, Wang L, Li H. Ad hoc-based feature selection and support vector machine classifier for intrusion detection, in 2007 IEEE International Conference on Grey Systems and Intelligent Services, 2007; pp. 1117-1121. https://doi.org/10.1109/GSIS.2007.4443446

[77] Berwick R. An Idiot's guide to Support vector machines (SVMs).

[78] Ahmad I, Abdulah AB, Alghamdi AS. Towards the designing of a robust intrusion detection system through an optimized advancement of neural networks, in Advances in Computer Science and Information Technology, ed: Springer, 2010; pp. 597-602.

[79] Han J, Kamber M, Pei J. Data mining: concepts and techniques: Elsevier 2011.

[80] SVM. Available: http://www.nickgillian.com/wiki/pmwiki.php/GRT/SVM

Received on 07-08-2017 Accepted on 11-08-2017 Published on 29-08-2017

https://doi.org/10.6000/1927-5129.2017.13.76