

A Survey on Incremental Learning

Junwei ZHONG^{1,a}, Zhenyan LIU^{1,b,*}, Yifei ZENG^{1,c}, Lijia CUI^{1,d} and Zizheng JI^{2,e}

¹Beijing Key Laboratory of Software Security Engineering Technology,
School of Software, Beijing Institute of Technology, Beijing 100081, China

²School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China

^a394134816@qq.com, ^bzhenyanliu@bit.edu.cn, ^c453624629@qq.com, ^d602465981@qq.com,
^ejzz@bit.edu.cn

Keywords: Classification; Sample Incremental Learning; Class Incremental Learning; Feature Incremental Learning

Abstract: Incremental learning is one of the research hotspots in machine learning. In this paper, we view the complex changes of data as three changes that are the change of sample, the change of class and the change of feature, and analyze the popular machine learning classification algorithms which support incremental learning. And then we focus on reviewing the research of three types of incremental learning: Sample Incremental Learning, Class Incremental Learning and Feature Incremental Learning. Finally, we make a prospect on the focus and difficulty of future research of incremental learning.

1. Introduction

One of the assumptions implied by the automatic classification of data is based on the statistical learning method is that the training data that used to train the classification model and the data that set to be classified come from the same population distribution, and the feature space of both is same(the feature space used by the data to be classified is constructed by the training stage by default). However, as time goes on, the data distribution and data features will change. After a while, the classification model that is trained based on historical data may not apply to some new data. Therefore, it is necessary for an automated classification system practical application oriented to achieve incremental learning and dynamically update, in order to ensure that the ever-changing data can be correctly classified.

The change of the data distribution can be further divided into the change of the sample and the change of the class. The so-called change in the sample refers to the change of the eigenvalues of the samples under the characteristic isomorphic space and the possible changes in the proportion of the samples in each class. The change of the class refers to the emergence of new classes, that is, the original classification system has changed. And the new data features may appear, that is, the feature space of the training data and the data that set to be classified is not same any more. In order to cope with these three different changes in the data, this paper divides the incremental learning approach into three categories: Sample Incremental Learning, Class Incremental Learning and Feature Incremental Learning. And the following carry out in-depth research and comprehensively review on these three kinds of incremental learning respectively.

The organizational structure of this paper is as follows: in the next four subsections, this paper firstly introduces the relevant background knowledge of incremental learning; then analyzes a variety of machine learning classification algorithms that can support incremental learning; and focuses on reviewing research progress of three kinds of incremental learning in the fourth section; and summarizes the full text and analyzes the further direction of incremental learning in the last section.

2. The background knowledge of incremental learning

In the traditional practical application environment, machine learning is a gradual process. Firstly, the training data is marked, and then the first training of the marked data is carried out to form a classification model that used to predict and classify data. When the new data need to be updated later, the previous training data set and the data set which need to be update dare merged into a new complete data set, and then the new complete data set is re-trained.

With the rapid development of Internet technology and the extensive application of computer technology, all kinds of industries gradually accumulate a large amount of data and still continue to increase in geometric multiple. The problems in the data analysis and in the data processing are as follows:

1. The acquisition of massive marked data requires a lot of manpower, material and time.
2. When the data set increases to a certain extent, the training time of the data set will be very long, resulting in time consumption and poor user experience.
3. When re-training the data set in traditional way, the program will be in a closed state, which is not suitable for online applications.
4. Most importantly, when a new class is generated, the general machine learning classification method can't automatically identify and update the new class. So it can't meet the actual needs.

When faced with the above problems, the general machine learning classification methods seem to stretch, or even powerless. Therefore, the concept of incremental learning was born. As early as 1962, both Coppock and Freund published article on incremental learning [1] in "Science". But today, there is still not an authoritative definition of incremental learning. Comprehending related literature [2], incremental learning should have four characteristics.

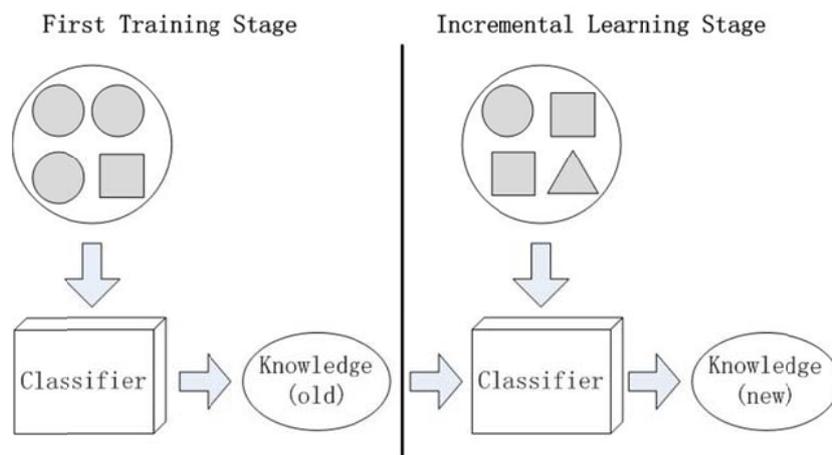


Figure1. Incremental learning process

According to Figure 1, the following describes four characteristics separately in the incremental learning process:

1. Classifier can get more information from the new data. As shown in figure 1, the classifier can learn the change of ratio of circles and squares from the new samples in incremental learning stage.

2. The data used for learning and training is only the new data. And there is no need to repeat to train the trained data. As shown in figure 1, there is no need to repeat to train the samples in the previous training stage. It is only need to train new samples in the incremental learning stage. Other than the traditional learning, which is necessary to integrate old and new samples into a complete set for re-training.

3. Acquiring new knowledge, it can integrate the old knowledge. That is to say knowledge learning process is a process of gradual accumulation. As shown in figure 1, when learning and training the new samples, it is need to establish on the previous knowledge system in the incremental learning process. It not only revises the previous knowledge system according to the new knowledge, but also integrates with new knowledge. And finally it form a new and more abundant knowledge system.

4. Classifier can adapt, identify and update new classes. As shown in figure 1, in the incremental learning stage, classifier learns the triangle which is a new class and never be trained before. So it can't be identified in traditional learning process. But incremental learning can identify the triangle, and expand the triangle into the collection of previous classes.

In the current study of incremental learning, some achievements have been made in the research direction of characteristic 1 (learning new knowledge), characteristic 2 (only training new data) and characteristic 3 (fusing new and old knowledge). However, research on characteristic 4 (identifying new classes) is still in an early stage, and few studies are relevant. At the same time, characteristics 3 (fusing new and old knowledge) refers to the concept of fusion, simple fusion means plus. In contrast, there is another concept of diminishment that means to decrease the old knowledge. That involves the concept of reduced learning [3], and it can also adjust the accuracy of the classification. The focus of this paper is the research of incremental learning, so the research of reduced learning will not take up space.

For the classification of incremental learning algorithm, it can be divided into three classes according to the difference of data change: Sample Incremental Learning (SIL), Class Incremental Learning (CIL) And Feature Incremental Learning (FIL). The following will introduce the above three classes in detail.

1. Sample Incremental Learning (SIL)

Firstly, a classification model is obtained based on the initial training data set in the training stage. The new large amount of data will constantly appear in the process of predicting the classification of the data. For a variety of reasons in the new data, the eigenvalues of the samples may change and the proportion of each class will also make a big change. Then it will affect the accuracy of classification. Therefore, it is need to ensure that the in case of existing knowledge(Generally, it will not change existing knowledge, but diminished learning need to modify existing knowledge), extract new knowledge through the incremental learning of new sample, fuse the old and new knowledge to improve the accuracy of classification. For example: in the recent network security trends, mobile Internet malicious programs, security vulnerabilities and other attacks show an upward trend, but botnet shows an downward trend compared to 2015.

2. Class Incremental Learning (CIL)

The initial training data set has a predetermined class system. In the process of predicting the classification of the data, a new category will be encountered, which is clearly different from the other class in the past. But bounded to belong to other classes, it will inevitably affect the overall performance of the classification. And the change of data isn't perceived, either. When new classes appear, the Class Incremental Learning needs to identify new classes and add them into the previous collection of class to improve the accuracy and intelligence of the classification. For example: in recent years, APT attacks will undoubtedly continue to lead the tide of Internet threats. While looking through the history of malicious attacks on the network, APT attacks appeared significantly later than other malicious attacks.

3. Feature Incremental Learning (FIL)

Through the training and learning of the initial training set, a data representation feature space is constructed in training stage. In the process of predicting the classification of the data, some new attribute features will be encountered that can promote the ability of the classification to a great extent and improve the accuracy of classification. Therefore, it is necessary to add new attribute features on the basis of existing feature space and construct new feature space to improve the accuracy of classification. For example: botnet continuously upgrades technical measures to prevent their malicious behaviors from being identifying. So it is bound to show new network behavior characteristics.

3. Classification algorithm that can support incremental learning

In the existing machine learning algorithm, there are many kinds of machine learning classification algorithms which support incremental learning (that is to say the classification principle of the classification algorithm has the characteristics that naturally support incremental

learning): Naive Bayesian, Support Vector Machine, Decision Tree, Random Forest, Artificial Neural Network, K-Nearest Neighbor, Fuzzy Rough Set Theory and Learn++NC, etc. This section mainly introduces Naive Bayesian, Support Vector Machine, Decision Tree, Random Forest and Artificial Neural Network.

3. 1 Naive Bayesian

Naive Bayesian algorithm [4] came from the 1950s and 1960s. The basic idea is to derive the posterior probability by calculating the prior probability and data. It has uncertain knowledge expression, rich probability expression ability and relatively strong learning and reasoning ability, which makes the study of Naive Bayesian enduring and become one of the best choice models in machine learning

The Naive Bayesian classification algorithm has the ability to support incremental learning is due to it has full advantage of the priori information. Posterior probability is predicted by prior probability and existing statistical data. This acquisition of the posterior probability will become the prior probability of next learning. Incremental learning uses this principle and uses the formula [5] to calculate new knowledge that is learn from new data based on old knowledge, in order to achieve the fusion of old and new knowledge. Ultimately, by adjusting to the prior probability, it improves its classification ability and accuracy.

3. 2 Support Vector Machine

Support Vector Machine (SVM) [6] was a machine learning method that firstly reported on Computational Learning Theory Conference in 1992, proposed by Corinna Cortes et al. It is divorced from the traditional method of thinking of reducing dimension, and used the reversal technique to increase the dimension of the problem space on purpose. Trying to find the best point between model complexity and learning ability based on limited sample information, it obtains the best reasoning ability and makes the classification problem easier.

Because of the concept of Support Vector (SV) makes the choice of the optimal hyperplane is determined by a small number of support vectors, Support Vector Machine classification algorithm has the ability to support incremental learning. Apply it to the incremental learning. When performing incremental learning, it learns new support vector in the incremental learning process based on previous support vector reserved before. It incorporates a few new support vectors that are important in new data and the support vectors reserved in the past, finally selects a few decisive support vectors to classify.

3. 3 Decision Tree

In 1975, J. Ross Quinlan introduced the concept of information entropy into the construction of Decision Tree, forming the famous ID3 algorithm [7]. It transforms a complex multi-category classification problem into a tree, starts from the root node, uses the hierarchical form from top to bottom, and selects the most powerful attribute through the information entropy to distribute the sample to the child node until the labeled leaf node (that is, all subsets on the leaf node contain only the same class of sample). C4.5 is an updated division tree, which introduces two spots as information gain ratio and pruning that can better extract feature for non-leaf node and solve the problem of data over fitting.

The Decision Tree has the ability to support incremental learning, because it is a tree structure and has inheritable characteristics. After learning new knowledge, ID4 algorithm [8] calculates information gain by the division of the attribution to determine whether to inherit or reconstruct the subtree. But the possibility of abandoning a lot of historical knowledge exists. Therefore, ID5R algorithm [9] can effectively inherit most of the historical calculation in the way of subtree promotion. Incremental learning of Decision Trees is undoubtedly the focus of the study of incremental learning.

3. 4 Random Forest

Random forest classification algorithm [10] is proposed by Leo Breiman and Adele Cutler et al

in the 1980s. Its main idea is integration thought, which is based on the development of decision trees, and includes multiple decision trees. The random forest integrates the results of all Decision Tree classification voting, specifying the class with the highest number of votes as the final output.

Random forest has the ability to support incremental learning, because it consists of multiple decision trees. It's not only has integrated idea, but also has some characteristics of the Decision Tree. In the incremental learning process, integrated idea makes it possible to achieve incremental learning by learning new knowledge and training a decision tree and then adding it into original random forest. And the Decision Tree also gives it the inheritance and other characteristics.

3.5 Artificial Neural Network

Artificial neural network (ANN) [11] originated in the 1940s, and proposed in a forward-looking article published by American mathematician Pitts and psychologist Mcculloch. They proposed MP model that uses neurons as logical operators to implement the algorithm. In 1957, Rosenblatt proposed a perceptual model that is the first genuine neural network model [12] based on MP. The model uses the knowledge of network topology as a theory to simulate the nervous system of the human brain and process the complex information inputted from the outside world. It has the characteristics of high intelligence, high fault tolerance and high parallelism.

Artificial neural network has the ability to learn incremental knowledge by itself, which is a nonlinear and self-adaptive information process system that composed of a large number of neurons interconnected. It can not only deal with data that happens a variety of changes, but also can handle each data in the way of incremental learning that affect the updating of the weight in some extent, and make their own power system constantly changing. The evolution of the entire dynamical system is a process of continuous iteration.

4. A review of incremental learning

This section will introduce the progress of research related to incremental learning in detail. It mainly elaborates from three aspects: Sample Incremental Learning, Class Incremental Learning and Feature Incremental Learning.

4.1 Sample Incremental Learning

Luo F X [5] proposed a new type of weighted Naive Bayesian incremental learning method, to solve the problem of incremental learning based on naive Bayesian algorithm in text field. There were three main improvements in its thinking. Firstly, it set the class confidence threshold, in order to strictly filter incremental samples. Secondly, according to the ability of classification, the coefficient was manually dynamically adjusted to improve the classification accuracy. Finally, using the word frequency weight formula which can reflect the importance of the feature, to overcome the defects can't highlight the different characteristics of the weight.

Ma H F et al. [13] proposed an incremental learning sequence selection method based on Naive Bayesian classification algorithm, to prevent the wrong classification from being added to the classifier and control the incremental training set scale. Firstly, it filtered the noise sample to select the correct sample to compose a new increment training subset. Then the new subset was filtered by using the 0-1 loss rate. By reasonably strengthening the learning sequence, preferentially choosing a sample with a strong class orientation, and considering the controversial noise samples, it woke its negative effects.

Pi W J et al. [14] proposed a data-driven parallel incremental Adaboost-SVM algorithm (PIASM) based on Hadoop, to solve the problem that Support Vector Machine (SVM) algorithm is difficult to adapt to large data. During the course of the study, they used the weight to characterize the distribution characteristics of each sample space, and then used the forgetting factor to adjust the size of the internal samples and quasi-boundary samples, to increase new samples which had greater impact and eliminate the historical samples which had less impact. PIASM improved the parallel processing capability of Support Vector Machine algorithms for large data, and also implemented incremental learning.

Ye Z X et al. [15] proposed an unbalanced SVM incremental learning method based on massive data, through extracting two aspects of features as URL-related features and web text content. In the study of online intelligent detection of phishing sites they found that new samples that satisfied the KKT condition couldn't change the support vector set, but new samples that violated the KKT condition could cause changes in the support vector support set. So the goal of incremental learning in experiments was to find new samples that could change the set of support vectors and train them.

Xie G D et al. [16] proposed an incremental learning method based on the K-Nearest Neighbor algorithm. By introducing the concept of layer to the model cluster, it established different layers in the model cluster to incremental data. The new model cluster could cover the original model cluster in the wrong area, giving a more accurate model cluster with higher layer values. When the new sample was covered by multiple layers, it selected the highest layer value of the model cluster as a classification of the results, and then achieved the goal of incremental learning.

Sun W X [17] proposed an incremental learning method based on Decision Tree in the study of fault diagnosis based on data mining and information fusion. It changed the single list in the original C4.5 Decision Tree to double-linked list, and added the function of remembering sample to the leaf node. New samples inherited the information of other branches when adding new samples. It achieved the incremental learning by judgmental forecasting whether the branch leaf nodes was necessary to subdivide and to modify the information of the branch.

Wang A P et al. [18] proposed an Incremental Extremely Random Forest (IERF) to solve the problem of online learning of small sample data streams. The new sample in the IERF algorithm was stored on the corresponding leaf node. And through the Gini coefficient, it determined whether to split the current leaf node. The experiment on UCI dataset proved that the performance of the incremental learning algorithm is better than the greedy decision tree reconstruction algorithm on moderate scale sample set.

4.2 Class Incremental Learning

Su Z T et al. [19] improved The Naive Bayesian classification model on two aspects, in order to solve two problems that there was no classification and update for new class in Naive Bayesian classifier and unbalance of the number of samples between classes would cause deviation to Naive Bayesian classification model. On one hand, it added parameter correction formula, so that it could adapt, identify and update new class. On the other hand, it introduced the Minimum Risk Decision Thinking to mitigate the effects of data imbalances.

Ye Q et al. [20] proposed a class incremental learning method based on the K-Nearest Neighbor algorithm. Firstly, it calculated the matching degree between the new sample and the standard sample, found the best matching sample and the secondary matching sample. Through comparing with the matching threshold, it determined to learn in classes or between classes. The algorithm could extend the knowledge structure of the original classifier, then identified and updated the new class.

Qin Y P et al. [21] proposed a new class of incremental learning methods based on Super-Ellipsoids. For each type of sample, they obtained a minimum super-ellipsoid that surrounds the class as much as possible in the feature space, so that the class could be distinguished by the super-ellipsoid. In the process of incremental learning, only samples in new class could be trained. By calculating whether the sample was included in the super-ellipsoid, it determined the class in classifying.

Xie M Q [22] proposed a class incremental learning method based on Decision Tree and Gaussian Mixture model. In the process of identifying new classes, it selected a sample set that was different from the existing category and was representative as the new class. Based on characteristics such as inheritance data mining ideas, it dug the similarity of two adjacent learning and the gradual change of knowledge, to extract the inheritable intermediate calculation for updating the parameters in the new class.

Dewan Md et al. [23] proposed an incremental learning method based on Decision Tree aim at the field of data mining. The process of adding new class or deleting to the old class was called

concept drift. When the decision tree constructed, the similarity degree of the leaf node attribute value was preliminarily clustered. If the number of data points classified by leaf nodes increased compared to the previous one, then checked whether it belonged to a cluster, to determine whether a new class appeared.

G Ditzler et al. [24] found that the Learn++NC algorithm had the ability to learn new class, and SMOTE had the ability to perform machine learning from unbalanced data sets. Combined with the characteristics of both, they proposed a new Learn++NC hybrid algorithm packaged by SMOTE. Even if the new class data was not balanced, this algorithm could still increase a new class or delete an old class. That is to say, it had a good plasticity.

H Jia et al. [25] proposed a new class incremental learning method based on Support Vector Machine (SVM) in the field of information retrieval and defect inspection. When new knowledge discovery process indicated that new data occupied the space outside the current knowledge domain in the feature space, it meant that these new data samples formed a new class, and used the new incremental learning algorithm to learn.

4.3 Feature Incremental Learning

Feng J [26] proposed an incremental learning method (HISVML) which could effectively solve the massive sample learning based on Support Vector Machines (SVM) in multi-layer text classification field. The feature incremental learning used batch learning to avoid the drift of concept. In the multi-layer classification concept tree, each node actually corresponded to an SVM classifier, and each non-leaf node maintained a feature set. The feature incremental updating was only related to the parent node and its ancestor nodes of the new sample set, and the other nodes will not updated.

Zeng S F et al. [27] proposed a feature incremental learning method based on Naive Bayesian. When selecting non-tagged incremental samples, it used double thresholds that were consisted of class confidence threshold and minimum posteriori probability to filter samples. When classifying the incremental samples, it would determine whether the processed samples included new feature that was not existed in the feature space. When identifying a new feature, it would add the new feature into the feature space, and update the class condition probability of the new feature.

X Liu et al. [28] proposed an algorithm of feature incremental learning based on least squares method Support Vector Machine (SVM). The algorithm had two excellent features. First, depending on the distribution of the samples, different kernel functions could be applied to previous features and new features. Second, the algorithm used the previous structural parameters only to train new features. It not only shorted the training time, but also reduced the storage space.

S Ozawa et al. [29] proposed a feature incremental learning method based on artificial neural network and Incremental Principal Component Analysis (IPCA) [32] in the field of artificial intelligence facial recognition. Since the IPCA updated the feature space step by step by rotating the eigen-axes and increasing the dimensions, the input of the neural classifier must also be changed. So they derived an approximation of the update formula for memory items that corresponded to representative training samples which stored in the long term memory (RAN-LTM). Combined with the new sample, it could effectively rebuilt and re-training to adapt to the evolution of feature space.

A Zeng P et al. [30] proposed a feature incremental learning method based on Fuzzy Rough Set (FRS). The mixed information system's data (HIS) often changed over time, and the data was varied (such as: real, Boolean and set values etc). To solve the problems, they proposed HD distance concept. Combing Gaussian kernel and Euclidean distance, it generated fuzzy information particles and constructed new fuzzy rough set. The feature increment could be selected and updated by analyzing the change in the attribute set.

E Swere et al. [31] proposed a feature incremental learning method based on Decision Tree, in order to solve the problem of how to learn quickly and adapt to emergency happened around in the field of artificial intelligence. They designed a frequency table that recorded the acquired feature vector and could record enough information about the C4.5 Decision Tree. In the learning stage, it could learn new features, add them to the frequency table in the form of eigenvectors, and generate

a new decision tree.

5. Summary

In this paper, the study of the current situation of incremental learning indicates that there is a considerable degree of progress in sample incremental learning. On one hand, it can achieve sample incremental learning by improving some machine learning algorithms which have the incremental learning characteristic to a certain extent. On the other hand, combining with clustering and classification technology, it can improve in the selection and handling of samples. However, researchers are still faced with many difficulties in class incremental learning and feature incremental learning. So how to enhance the performance of class incremental learning and of feature incremental learning will be the focus of the next step study. In addition, how to integrate simple incremental learning, class incremental learning and feature incremental learning together organically to cope with complex data changes, that is also a research direction worth to further study.

Acknowledgements

This work was financially supported by Frontier and Interdisciplinary Innovation Program of Beijing Institute of Technology (NO.3080012291701).

References

- [1] H.W.Coppock, J.E.Freund. All-or-none versus incremental learning of errorless shock escapes by the rat [J]. Science, 1962,135(3500):318-319.
- [2] POLIKAR R, UPDAL, UPDASS, et al. Learn++: An incremental learning algorithm for supervised neural networks [J]. IEEE transactions on systems, man, and cybernetics, part C (Applications and reviews), 2001,31(4):497-508.
- [3] Duan H, Hou W Z, He G P, Lian W J. Incremental and Decremental Learning with Support Vector Machine [J]. Journal of Harbin Engineering University, 2006,27(b07):415-421.
- [4] CHRISTOPHER M. Pattern recognition and machine learning [M]. Springer, 2006.
- [5] Luo F X. Research on Incremental Learning Naive Bayesian Chinese Classification System [D]. 2008.
- [6] Vapnik V N. The nature of statistical learning theory [N]. Springer-Verlag, 1995.
- [7] Tom M. Machine Learning [M]. 2007:40.
- [8] JC Schlimmer, DH Fisher. A Case Study of Incremental Concept Induction [C]. National Conference on Artificial Intelligence Philadelphia, 1986,abs/0904,3352:496-501.
- [9] UTGOFF P E. Incremental induction of decision trees [M]. Machine Learning, 1989,4(2):161-186.
- [10] R Genuer, JM Poggi, C Tuleau-Malot, N Villa-Vialaneix. Random Forests for Big Data [J]. Computer Science, 2015.
- [11] W Pitts. A logical calculus of the ideas immanent in nervous activity[J]. MIT Press, 1988,5(4):115-133.
- [12] F Rosenblatt. The Perceptron - a perceiving and recognizing automaton [R]. Math Stat, 1957,85-460-1.
- [13] Ma H F, Fan X H. Improved incremental Bayes classification algorithm [J]. Chinese Journal of Scientific Instrument, 2007,28(8).

- [14] Pi W J, Gong X J. Data driven parallel incremental support vector machine learning algorithm based on Hadoop framework [J]. Journal of Computer Applications, 2016,36(11):3044-3049.
- [15] Ye Z X, Wang D H. Detection Method of Phishing Web Site Based on Massive Data Unbalanced SVM Incremental Learning [J]. Telecommunication Engineering Technology and Standardization, 2016,29(12):26-31.
- [16] Xie G D, Huang J, Chen L F. KNN Model Based Incremental Learning Algorithm [J]. PR & AI, 2010,23(5):701-707.
- [17] Sun W X. Research on Approaches of Data Mining and Information Fusion based Fault Diagnosis[D]. 2007.
- [18] Wang A P, Wan G W, Cheng Z Q, Li S K. Incremental Learning Extremely Random Forest Classifier for Online Learning [J]. Journal of Software, 2011,22(9):2059-2074.
- [19] Su Z T, Li Y. On Improved Incremental Bayesian Classification Model [J]. Computer Applications and Software, 2016,33(8):254-259.
- [20] Ye Q, Lu Z H, Zhou J, Song Z. Research on incremental learning classification algorithm based on near neighbor [J]. Computer Engineering and Applications, 2016,52(20):154-157.
- [21] Qin Y P, Chen Y D, Wang C L, Wang X K. New class incremental learning method. Computer Engineering and Applications [J], 2011,47(34):181-183.
- [22] Xie M Q. Research on New Class Recognition and Classifier Updating Method in Classification Problem [D]. 2006.
- [23] Dewan Md, Farid, Chowdhury Mofizur Rahman. Novel Class Detection in Concept-Drifting Data Stream Mining Employing Decision Tree [J]. International Conference on Electrical and Computer Engineering, 2012:630-633.
- [24] G Ditzler, G Rosen, R Polikar. Incremental learning of new classes from unbalanced data [C]. Proceedings of International Joint Conference on Neural Networks, 2013.
- [25] H Jia, YL Murphey, D Gutchess, TS Chang. Identifying knowledge domain and incremental new class learning in SVM [J]. IEEE International Joint Conference on Neural Networks, 2005,5(5):2742-2747.
- [26] Feng J. Research on Multi - layer Text Classification and Incremental Learning Key Technology [D]. 2009.
- [27] Zeng S F, Zhang X Y, Du X F, Lu T B. Improved incremental algorithm of Naive Bayes [J]. Journal on Communications, 2016,37(10):81-91.
- [28] X Liu, G Zhang, Y Zhan, E Zhu. An Incremental Feature Learning Algorithm Based on Least Square Support Vector Machine [J]. Springer-Verlag Berlin Heidelberg, 2008,5059(12):330-338.
- [29] S Ozawa, SL Toh, S Abe, S Pang, N Kasabov. Incremental learning of feature space and classifier for face recognition [J]. Neural Networks, 2005,18(5-6):575-584.
- [30] A Zeng, T Li, D Liu, J Zhang, H Chen. A fuzzy rough set approach for incremental feature selection on hybrid information systems [J]. Fuzzy Sets & Systems, 2015,258(1):39-60.
- [31] E Swere, D Mulvaney, I Sillitoe. A Fast Memory-Efficient Incremental Decision Tree Algorithm in its Application to Mobile Robot Navigation [J]. IEEE/RSJ International Conference on Intelligent Robots & Systems, 2006:645-650.