

A review on data stream classification approaches

Sajad Homayoun *, Marzieh Ahmadzadeh

Department of Computer Engineering and Information Technology, Shiraz University of Technology, Shiraz, Iran

*Corresponding author E-mail: s.homayoun@sutech.ac.ir

Abstract

Stream data is usually in vast volume, changing dynamically, possibly infinite, and containing multi-dimensional features. The attention towards data stream mining is increasing as regards to its presence in wide range of real-world applications, such as e-commerce, banking, sensor data and telecommunication records. Similar to data mining, data stream mining includes classification, clustering, frequent pattern mining etc. techniques; the special focus of this paper is on classification methods invented to handle data streams. Early methods of data stream classification needed all instances to be labeled for creating classifier models, but there are some methods (Semi-Supervised Learning and Active Learning) in which unlabeled data is employed as well as labeled data. In this paper, by focusing on ensemble methods, semi-supervised and active learning, a review on some state of the art researches is given.

Keywords: Data Stream; Data Stream Classification; Ensemble; Semi-Supervised Learning; Active Learning.

1. Introduction

Dramatic growth in information technology and vast volume of generated data has made new challenging discovery tasks in processing of data. The term "data stream" is defined as a sequence of data that arrives at a system in a continuous and changing manner. Data stream can be conceived as a continuous and changing sequence of data that continuously arrives at a system to be stored or processed [1]. Data streams have some characteristics in common such as massive, temporally ordered, fast-changing and potentially infinite in length [2-4]. According to [5], there are some reasons which dispart data streams from traditional data mining:

- The size of data streams is potentially boundless.
- The elements of stream arrive on-line.
- Because of limitations in memory space, after processing of an element, system discards (or summarizes) it.
- The system cannot control or determine how data elements arrive.

Emails, sensor data, websites customer click stream, network traffic, weather forecasting data etc. are some examples of data stream. Data stream mining comprises three main techniques such as clustering, classification and frequent pattern mining.

Classification is a supervised learning techniques which aims to predict of an independent variable (class label) according to some values of an instance[6]. Making a classification model has two main phases: 1) Model creation, 2) Model evaluation. At the first phase, a learning algorithm uses dataset to create a model which is able to predict class label. The second phase tries to investigate the accuracy parameters of created model.

Change in data according to time is one of the main issues in data stream classification techniques. There are two evolution in data [7]: 1) concept drifting, 2) concept evolution. Concept drifting happens whenever class labels changes due to changes in time. Weather forecasting, spam categorization and monitoring systems are some examples in which concept drifting is a challenge. Concept evolution occurs when one or more new class labels emerge

on class label set [7]. As shown in Fig. 1 b, concept evolution occurs when new instances arrive with new class labels.

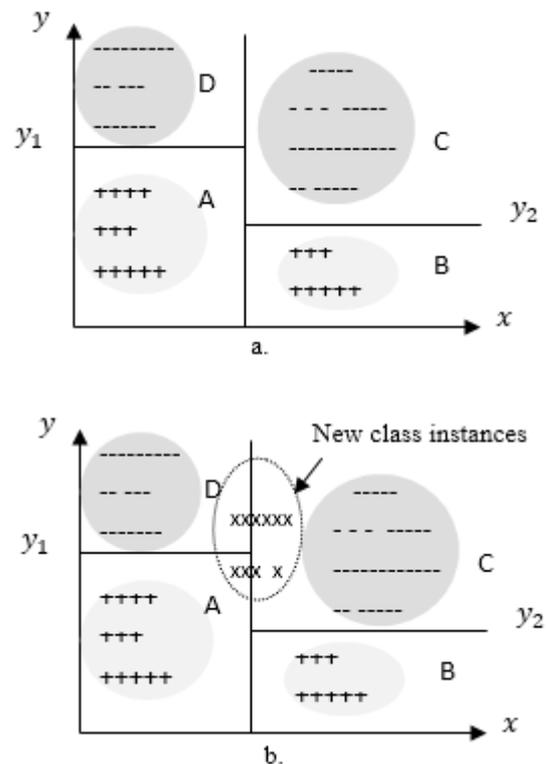


Fig. 1: (A) Fixed Number of Class Labels; (B) A Novel Class Has Emerged (Concept Evolution)

Classification in data stream has some challenges that researchers attempt to solve them. Three main challenges of classification techniques are as follows [8]:

- Accuracy: It is the most important factor in classification algorithms, and concept drifting directly influences the accuracy.
- Efficiency: creating of a classifier is costly from processing point of view. Also, updating of the model is a challenge due to drifting.
- Ease of use: a classifier model should be usable in applications.

According to [9], single model incremental and ensemble-based classification are two major branches of data stream classification. The first works on one single classifier and update it incrementally to tackle new evolved stream class labels. It usually needs complex modifications on the internal structure of the classifier. Single classifier approach often is unable to create strong and accurate classifiers. In contrast, an ensemble model combines different classifiers to improve the overall accuracy of predictions. If every single classifier works better than random prediction (accuracy more than 0.5), then ensemble model is always more accurate than a single classifier model.

Due to the need for labeled instances to build classifiers, researchers contemplate classification as supervised. It's worth mentioning that quality of classifiers extremely depends on the percentage of labeled instances available in data stream. Many researchers have tried to use unlabeled instances as well as labeled instances be-

cause manual labeling of instances (by experienced agents) is costly and time consuming.

The reminder of the paper is organized as follow: section 2 discusses some ensemble classification methods; section 3 presents a review on semi-supervised and active learning algorithms while section 4 is dedicated to future works, and finally section 5 concludes the paper.

2. Incremental learning and ensemble methods

After introducing ensembles in 1990s, many researchers tried to improve prediction accuracy by using ensembles [10]. An ensemble method(Fig. 2) creates a set of base classifiers from training data and classify new instances by pooling of base ensembles [11]. Ensembles are popular because they improve classification accuracy in static environments [12]. But they need some changes to adapt with dynamic nature of data streams. In Incremental learning, a machine learning algorithm take place when new instances emerge, and then to adjust the model [13]. Some methods of reviewed incremental methods are suggested in section 3 because they are in the category of Semi-Supervised or Active Learning algorithms. In this section, a review on some incremental and ensemble methods is given.

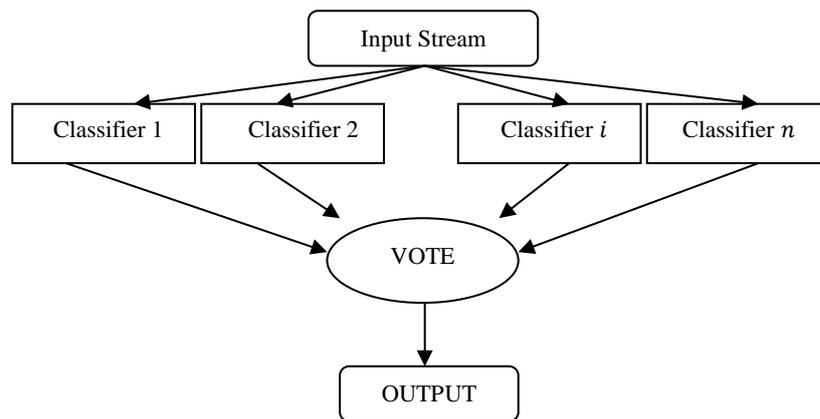


Fig. 2: An Ensemble Model.

Jing et al. [14] introduced four main challenges of classification techniques and claimed their proposed model is able to address all challenges: 1) infinite length, 2) concept drifting, 3) arrival of novel classes and 4) lack of labeled instances. For handling concept evolution and lacking of labeled instances, a novel class detection mechanism is proposed. ECM-BDF (Ensemble Classification Model Based on Decision Feedback) divides data streams into sequential chunks with appropriate sizes, then a classifier is made for each data chunk and some of created classifier considered as ensemble. In addition, the classifiers made from new labeled instances used for updating of the ensemble. There is also a novel class detection mechanism to face arriving new class labels and this mechanism assumes a decision boundary around training data. The data which place out of boundary considered as outliers. And, Outliers with strong cohesion may consider as arriving new class. The proposed model in [14] only uses labeled instances, but in fact in some cases unlabeled instances are much more than labeled ones and models which only consider labeled instances usually have low accuracy. Abdulsalam et al. [15] defined four scenario for data stream and presented a three phase model which addresses the scenarios. Scenario 0 in which labeled records only appears at the beginning of the stream (and is enough for creating classifier) and the consequent instances are unlabeled. Scenario 1 shows concept drifting while labeled instances are adequate for making classifier. Scenario 2 and 3 in which there are no sufficient labeled instances. Scenario 3 is more common and shows arriving of la-

beled record frequently and periodically. Fig. 3 shows four mentioned scenarios.

In phase one, they introduced an approach to handle scenario 0 in which stream decision tree construction is merged with Random Forest algorithm. Phase two uses self-adjusting algorithm which employs entropy-based change-detection technique to address scenario 1(concept drifting). Phase three aims to handle scenario 2 and scenario 3 while the key feature of this phase is the ability of determining when the current model is ready to deploy. In other words, it determines deployment moment by defining a threshold value for minimum number of needed labeled records. Proposed algorithm in [15] considers only ordinal or numerical attributes and it also assumes records are approximately uniformly distributed; these issues cause limitations on proposed model.

AUE2(Accuracy Updated Ensemble) is proposed in [16] aims to handle different types of drifts. It combines accuracy based weighting mechanisms achieved from changes in block based ensembles with nature of Hoeffding Trees. Ensemble is updated with appending new classifiers and removing weak classifiers. In other words, the proposed model improves ensemble reactions when facing different drifts while decreasing influences of data chunk size on prediction accuracy. AUE2 is an enhance for AUE1 [17] with some changes in weighting and updating mechanisms to reduce computation cost and to increase accuracy of prediction.

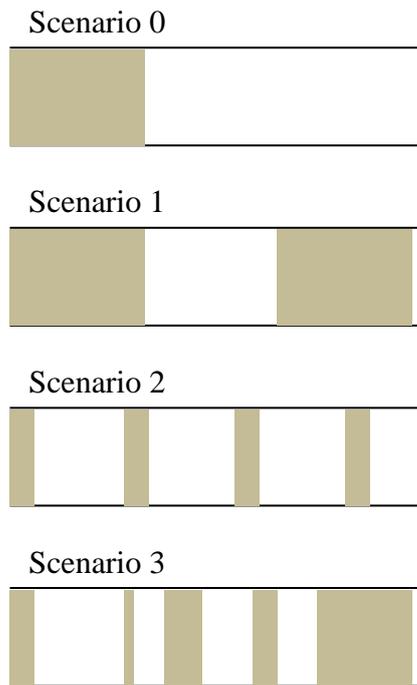


Fig.3: Scenarios Introduced in [15] for Data Streams.

[9] proposed an adaptive ensemble approach for classification and novel class detection in concept drifting. It uses traditional classifiers and applies automatic updating of ensemble models for handling concept drifting. The idea for novel class detection is the distance among instances. In other words, instances of a class should be closer and instances of different classes should be far enough. If an instance is apart enough from available clusters, it can consider as new class label.

Some articles tried to address feature-evolution; it occurs when set of features changes during time [18] or whenever new features appear in data[19].

DXMiner from [19] tried to introduce a model for handling of feature evolution, but it had high false positive rate (false novel class detection rate) and false negative rate (missed novel class detection rate) in some datasets. Moreover, it is unable to detect new classes if more than one new class arrives at a time. Afterwards, Masud et al. [20] tried to solve the problem of simultaneous arriving of new classes. A model is introduced in [21] which the authors claimed that it has more performance compared to earlier models in concept drift, concept evolution and feature evolution. To handle concept drift and concept evolution they designed a framework in which each classifier is equipped with a novel class detector and is able to detect more than one novel classes. To address feature evolution, they proposed feature set homogenization technique.

Aggarwal et al. proposed a model which is able to adapt with changes in data streams(concept evolution) [22]. They proposed On Demand Classification which is able to dynamically determine appropriate window size for past training data. They introduced supervised micro clustering which only made from training data and each micro cluster is a set of related training instances in which a cluster's instances have same class label. They tried to change unsupervised clustering approach [23] and handle high evolving data stream. Note that some parameters (such as initial points, size of sliding window etc.) must be set carefully to achieve appropriate accuracy and it seems as a drawback of their model. Their model aims at handling concept drifting while it has no idea for concept evolution. They used KDD 99 data set [24] to investigate the model.

3. Semi-supervised learning and active learning

Due to the nature of data stream (high volume, quick et.), labeling of instances (by experienced agents) is not possible and researchers tried to propose novel methods to handle this problem. According to [25], Semi-Supervised Learning(SSL) and Active Learning(AL) are two iterative approaches of employing unlabeled data in creating classification model.

By the purpose of reducing manual labelling workload, Semi-Supervised Learning aims to label samples by the machine itself, while Active Learning attempts to find the most informative samples for labelling by experts. The primary characteristics of SSL and AL are as follows:

- SSL: it selects the sample that has the highest confidence, and adds the predicted label by the machine itself without any external (expert) involvement at each iteration.
- AL: it takes the instance which has the lowest confidence as the most informative one; it selects such instance and asks the expert for its label in each iteration. AL involves human experts and aims at selecting the most useful instances for training. It can greatly improve the model's performance and can accelerate the speed of convergence.

Masud et al. [7] tried to employ unlabeled data in building classifier model as well as labeled data. Using the fact that the high percentage of data are unlabeled (because the speed of labeling of instances by experienced agents is less than the speed of arriving data, and earlier classification models only employed labeled data). For making prediction more accurate, their model tries to create classification model by using of both labeled and unlabeled instances. They introduced a semi-supervised clustering algorithm and build classifiers on evolving data by a label propagation approach. The model considers both challenges of concept drift and concept evolution. They compared their proposed model with On-Demand method proposed in [22] and the results shows it works better (in memory usage, computation time and accuracy) while it only uses 10 percent of training labeled data in compare On-Demand which uses 100 percent labeled data to build classifier.

Semi-Supervised Classification based on Class Membership (SSCCM) is proposed in [26] and uses label membership in semi-supervised learning. They formulated the problem for labeled and unlabeled data in a unified objective function. Afterwards, they solved it by using of an iterative strategy which tries to converge to final solution in each iteration. SSCM uses both label membership and decision functions for classification and prediction of functions are consistent. In other words, it is assumed an instance is near the decision boundary if two predictions are inconsistent and probably the prediction is unreliable. In fact one function is sufficient for prediction and label membership is preferred. Note that one can use inconsistency between two predictions to identify instances which are difficult to classify and use other ways of classification (such as manual labeling etc.). In particular, each function is verified by the other and the reliability of classification is improved.

SUN is a Supervised classification algorithm for data streams with concept drifts and UNlabeled data [27] aims to handle concept drifting with data streams including unlabeled data. SUN uses of a k-modes based algorithm which incrementally places concept clusters in leaves of constructed decision tree. Converting categorical data into numerical does not make meaningful results necessarily if there is no particular order in categorical data (traditional clustering algorithms). Therefore, the results of k-means and k-median are not appropriate and a k-modes based algorithm is introduced in [27].

According to the theory of Naïve Bayes, for a fixed (and unchanging) distribution of the instances, the online error of Naive Bayes will decrease; while the online error of Naive Bayes will increase for changing instances. In [28], the change in data distribution demonstrates the change in attribute dimensions. Thus, to deal with concept drifts, SUN compares the history concept to new

concept and considers the distribution of class label to track concept drifts.

Zhang et al. proposed an ensemble model in [28] which uses a combination of classification and clustering for mining data streams. They introduced two challenges for combining of two mentioned methods: 1) generated clusters having only a cluster number and there is no information about instances of a cluster, 2) due to concept drifting, combining of clusters and classifiers in one ensemble is a difficult task. Zhang et al. proposed a solution for handling of each mentioned challenge: 1) using of a label propagation technique for each cluster to extract useful information (label) from instances of a cluster, 2) weighting approach to weigh classifier models based on consistency to constructed model from up to date data chunks. [28] assumes available class labels of unlabeled data chunks are similar to labeled chunks and it means there is no solution to handle concept evolution.

A classifier ensemble-based active learning framework is proposed in [28] which selectively labels instances to build an ensemble classifier. [28] proves classifier ensemble's variance directly adapt error rate; and classifier ensemble's variance is equal to the accuracy of prediction. Hence, agent should label instances to minimize classifier ensemble variance and Minimum Variance (MV) is proposed. To determine weight values for ensemble classifier, an optimal weight calculation method is proposed in [28]. Finally, MV and optimal weighting is combined to make a framework.

Hosseini et al. in [29] tried to make use of recurring concept in learning data stream classifier. They used two approaches of Active Learning (AL) and weighted classifier. There is a pool of classifier which be updated continuously and each classifier in the pool describes one of the existing concepts. When new data arrives, the model classifies the instances and after determining the

label, an existing classifier available in the pool updated or a new classifier inserted into the pool. Two methods of Bayesian and heuristic are used for detecting of recurring concepts and updating the pool.

4. Future works

AL can help in cases in which there is an expert to determine class labels and track the model toward high accuracy. In cases of arriving vast volume of data in extremely high speed, it may output low accuracy and using of semi-supervised is preferred. SSL tries to automatically find useful information from unlabeled data (and it means SSL is high in speed), but in cases that the initial model is very weak, it might produce wrong labels and cause mistakes in training set. Furthermore, the instances having the highest confidence are not necessarily the most useful ones, so SSL generally performs worse than AL. Combining of AL and SSL seems as a research gap.

Concept evolution (especially arriving new classes simultaneously) needs more attempt because few researches are available at this area and the proposed methods are often too complicated.

An integrated ensemble model in which all challenges such as concept drifting, concept evolution and scarcity of labeled instances is needed. However, there are some researches on the topic, but proposed models usually tested on well-known datasets and systems experiments and implementations in longer period of time are needed.

Table 1 shows a brief on investigated researches and it can help readers to find topics for future works.

Table 1: A Comparison of Investigated Research Papers

Author(s)	Year	Short Description	Case	CD?	CE?	Tech.	Approach
Masud et al. [7]	2011	Utilizing both labeled and unlabeled instances to train and update classification model	SynD, SynDE, KDD 99', ASRS	Y	Y	E	Semi-supervised clustering + Label propagation
Aggarwal CC et al. [22]	2006	Considering only labeled instances of data and Building the classifier through an on-demand classification process which can dynamically select the appropriate window of past training data.	KDD 99'	Y	N	I	Supervised micro-clustering, Cluster-based, Sliding window
LIU Jing et al. [14]	2014	Data streams classification with ensemble model based on decision-feedback	SynCN, KDD 99'	Y	Y	E	Novel class label detection, feedback from unsupervised mechanisms
Masud et al. [19]	2010	Considering dynamic feature space and classification and addressing feature-evolution	Twitter, ASRS, KDD 99', Forest	Y	Y	I	Semi-Supervised, Lossless Homogenizing Conversion for feature-evolution
Xindong Wua et al. [27]	2012	Handling both challenges of concept drifting and unlabeled data streams.	SEA, STAGGER, KDD '99, Yahoo shopping data, LED.	Y	N	I	Semi-supervised, k-modes based clustering, statistical approach in detecting concept drifts
Peng Zhang et al. [30]	2010	Accumulating labeled records and combine them to create a classifier according to threshold.	The Malicious URLs Detection dataset. The Intrusion Detection dataset.	Y	N	E	Semi-supervised, Label propagation in clusters and weighting in updating ensemble framework.
Abdulsalam et al. [15]	2011	Considering multiple target class labels.	Synthetic dataset, Sloan Digital Sky Survey (SDSS)	Y	N	E	entropy-based concept drift detection, Random Forest
Dariusz Brzezinski et al. [16]	2014	Reacting to different types of concept drift.	Some datasets from UCI repository[31].	Y	N	E	Accuracy-based weighting, Hoeffding Trees

Xingquan Zhu et al. [28]	2010	Selecting best instances to determine labels by foreign agent by the purpose of decreasing classifier ensemble variance.	Data stream generated by Hyperplane-based synthetic data stream generator.	Y	N	E	Minimum Variance (MV), optimal weighting
Hosseini et al. [29]	2011	Construct a pool of classifier and updating the pool according to new classifier created from new arrived data stream to improve accuracy of the ensemble.	Data stream generated by Hyperplane-based synthetic data stream generator, and Emailing list dataset.	Y	N	E	Bayesian formulation, heuristic methods
Dewan Md. Farid et al. [9]	2013	Handling concept evolution by considering inter-class distance and intra-class distance.	Some datasets from UCI repository [30].	Y	Y	E	Decision Tree Learning, Similarity Based Clustering,
Yunyun Wang et al. [26]	2012	Enhancing classification reliability by consistency check between predictions of two functions. Each instance has likelihood to class labels instead of belonging to only one class.	Some datasets from UCI repository[31].	N	N	I	Semi-supervised, Label membership function, decision function

CD: Concept Drift, CE: Concept Evolving, E: Ensemble, I: Incremental, Y: Yes, N: No, Tech: Technique.

5. Conclusion

Data stream mining includes techniques such as classification, clustering, frequent pattern mining etc. In this paper, a review is given by focusing on ensemble methods, semi-supervised and active learning methods. Traditional data stream classification methods only employ labeled data and often have less accuracy in cases that system faces the lack of enough labeled data. In real-world, scarcity of labeled instances is usual because labeling is a time consuming and costly process. Hence, recently researchers have focused on using unlabeled data as well as labeled data in creating classification models. AL and SSL are two approaches of using unlabeled data. Handling concept drifting is an important issue on data stream mining, though few references focused on concept evolution as well.

References

- [1] H. H. Mahnoosh Kholghi, Mohammad Reza Keyvanpour, "Classification and Evaluation of Data Mining Techniques for Data Stream Requirements," presented at the International Symposium on Computer, Communication, Control and Automation, Tainan, Taiwan, 2010, <http://dx.doi.org/10.1109/3CA.2010.5533759>.
- [2] *Data Streams Models and Algorithms*: Springer, 2007,
- [3] J. D. U. Anand Rajaraman, *Mining of Massive Datasets*: Cambridge, 2012,
- [4] J. G. M. M. Gaber, *Learning from Data Streams*: Springer, 2007,
- [5] J. Gama, *Knowledge Discovery from Data Streams*: Chapman & Hall/CRC, Taylor & Francis Group, 2010, <http://dx.doi.org/10.1201/EBK1439826119>.
- [6] M. Kantardzic, *Data mining : concepts, models, methods and algorithms*: Wiley-IEEE Press, 2011, <http://dx.doi.org/10.1002/9781118029145.ch1>.
- [7] C. W. Mohammad M. Masud, Jing Gao, Latifur Khan, Jiawei Han, Kevin W. Hamlen & Nikunj C. Oza, "Facing the reality of data stream classification: coping with scarcity of labeled data," *Knowl Inf Syst*, vol. 33, p. 32, 2011, <http://dx.doi.org/10.1007/s10115-011-0447-8>.
- [8] O. M. L. Rokach, *Data Mining and Knowledge Discovery Handbook*, 2 ed.: Springer, 2010, <http://dx.doi.org/10.1007/978-0-387-09823-4>.
- [9] L. Z. Dewan Md. Farid, Alamgir Hossain, Chowdhury Mofizur Rahman, Rebecca Strachan, Graham Sexton, Keshav Dahal, Dewan Md. Farid, Li Zhang, Alamgir Hossain, Chowdhury Mofizur Rahman, Rebecca Strachan, Graham Sexton & Keshav Dahal, "An adaptive ensemble classifier for mining concept drifting data streams," *Expert Systems with Applications*, vol. 40, p. 12, 2013,
- [10] R. Polikar, "Ensemble based systems in decision making," *IEEE Circuits and Systems Magazine*, vol. 6, p. 25, 2006, <http://dx.doi.org/10.1109/MCAS.2006.1688199>.
- [11] M. S. V. K. Pang-Ning Tan, *Introduction to Data Mining* vol. 1: Pearson Education, 2006,
- [12] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*. USA: Wiley, 2004, <http://dx.doi.org/10.1002/9781118914564.refs>.
- [13] P. Z. Wenyu Zang, Chuan Zhou & Li Guo, "Comparative study between incremental and ensemble learning on data streams: Case study," *Journal of Big Data*, vol. 1, p. 16, 2014, <http://dx.doi.org/10.1186/2196-1115-1-5>.
- [14] X. G.-s. LIU Jing, ZHENG Shi-hui, XIAO Da & GU Li-ze, "Data streams classification with ensemble model based on decision feedback," *The Journal of China Universities of Posts and Telecommunications*, vol. 21, p. 7, 2014, [http://dx.doi.org/10.1016/S1005-8885\(14\)60272-7](http://dx.doi.org/10.1016/S1005-8885(14)60272-7).
- [15] D. B. S. P. M. Hanady Abdulsalam, "Classification Using Streaming Random Forests," *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, vol. 23, p. 15, 2011, <http://dx.doi.org/10.1109/TKDE.2010.36>.
- [16] D. B. J. Stefanowski, "Reacting to Different Types of Concept Drift: The Accuracy Updated Ensemble Algorithm," *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, vol. 25, p. 14, 2014, <http://dx.doi.org/10.1109/TNNLS.2013.2251352>.
- [17] D. B. J. Stefanowski, "Accuracy Updated Ensemble for Data Streams with Concept Drift," presented at the International Conference on Hybrid Artificial Intelligent Systems, 2011, http://dx.doi.org/10.1007/978-3-642-21222-2_19.
- [18] J. Gao. (2014-07-01). *Data Stream Mining: Challenges and Techniques*. Available: <http://www.cse.buffalo.edu/~jing/talks.htm>
- [19] Q. C. Mohammad M. Masud, Jing Gao, Latifur Khan, Jiawei Han & Bhavani Thuraisingham, "Classification and Novel Class Detection of Data Streams in a Dynamic Feature Space," presented at the European conference on Machine learning and knowledge discovery in databases, Berlin, 2010, http://dx.doi.org/10.1007/978-3-642-15883-4_22.
- [20] Q. C. Mohammad M. Masud, Latifur Khan, Charu Aggarwal, Jing Gao, Jiawei Han & Bhavani Thuraisingham, "Addressing Concept-

- Evolution in Concept-Drifting Data Streams," presented at the IEEE International Conference on Data Mining, 2010, <http://dx.doi.org/10.1109/ICDM.2010.160>.
- [21] C. W. Mohammad M. Masud, Jing Gao, Latifur Khan, Jiawei Han, Kevin W. Hamlen & Nikunj C. Oza, "Classification and Adaptive Novel Class Detection of Feature-Evolving Data Streams," *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, vol. 25, p. 14, 2013, <http://dx.doi.org/10.1109/TKDE.2012.109>.
- [22] J. H. Charu C. Aggarwal, Jianyong Wang & Philip S. Yu, "A Framework for On-Demand Classification of Evolving Data Streams," *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, vol. 18, p. 13, 2006, <http://dx.doi.org/10.1109/TKDE.2006.69>.
- [23] J. H. P. S. Y. Charu C. Aggarwal, "A Framework for Clustering Evolving Data Streams," in *International Conferences of Very Large Data Bases*, Berlin, 2003, p. 11,
- [24] (1999, 2014-06-06). *KDD Cup 1999 Data*. Available: <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>.
- [25] X. X. G. Q. Yan Leng, "Combining Active Learning and Semi-supervised Learning to Construct SVM Classifier," *Knowledge-Based Systems*, vol. in press, p. 31, 2014, <http://dx.doi.org/10.1016/j.knosys.2013.01.032>.
- [26] S. C. Z.-H. Z. Yunyun Wang, "New Semi-Supervised Classification Method Based on Modified Cluster Assumption," *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, vol. 23, p. 14, 2012, <http://dx.doi.org/10.1109/TNNLS.2012.2186825>.
- [27] P. L. X. H. Xindong Wua, "Learning from concept drifting data streams with unlabeled data," *Neurocomputing*, vol. 92, p. 11, 2012, <http://dx.doi.org/10.1016/j.neucom.2011.08.041>.
- [28] P. Z. Xingquan Zhu, Xiaodong Lin & Yong Shi, "Active Learning From Stream Data Using Optimal Weight Classifier Ensemble," *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART B: CYBERNETICS*, vol. 40, p. 15, 2010, <http://dx.doi.org/10.1109/TSMCB.2010.2042445>.
- [29] Z. A. H. B. Mohammad Javad Hosseini, "Pool and Accuracy Based Stream Classification: A new ensemble algorithm on data stream classification using recurring concept detection," presented at the 11th IEEE International Conference on Data Mining Workshops, 2011, <http://dx.doi.org/10.1109/ICDMW.2011.137>.
- [30] X. Z. Peng Zhang, Jianlong Tan & Li Guo, "Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams," presented at the 2010 IEEE 10th International Conference on Data Mining (ICDM), Sydney, NSW, 2010, <http://dx.doi.org/10.1109/ICDM.2010.125>.
- [31] (2014-07-08). *UC Irvine Machine Learning Repository*. Available: <http://archive.ics.uci.edu/ml/>.