Online Approach to Handle Concept Drifting Data Streams using Diversity

Parneeta Sidhu and Mohinder Bhatia
Division of CoE, Netaji Subhas Institute of Technology, University of Delhi, India

Abstract: *Concept drift is the trend observed in almost all real time applications. Many online and offline algorithms were developed in the past to analyze this drift and train our algorithms. Different levels of diversity are required before and after a drift to get the best generalization accuracy. In our paper, we present a new online approach Extended Dynamic Weighted Majority with diversity (EDWM) to handle various types of drifts from slow gradual to abrupt drifts. Our approach is based on the Weighted Majority(WM) vote of the ensembles containing different diversity levels. Experiments on the various artificial and real datasets proved that our proposed ensemble approach learns drifting concepts better than the existing online approaches in a resource constrained environment.*

Keywords: *Online learning, ensemble, concept drift, data streams, diversity*.

Received November 29, 2013; accepted December 16, 2014

1. Introduction

Online learning has been proved to be really important for handling real time application where data is arriving continuously with varying data distributions. The concept underlying the data is changing with time and we need to handle these data distributions within real time and space.

A concept refers to the distribution of training data being featured by the joint distribution [1], p(x, y) where x represents the n-dimensional feature vector($x=\{x_i\}$, $1\leq i \leq n$) and y represents the corresponding class label. So the concept drift means a change in the data distribution [6, 7], with the arrival of new data examples of such applications are market-basket analysis, credit card fraud detection, internet data, process control, intrusion detection [14] etc.

Online learning algorithms [5] take as input a single labelled training example as well as a hypothesis at each time step and output an updated hypothesis. Thus, for a sequence of training examples an online algorithm will produce a sequence of hypotheses. These approaches can be categorized as:

- Approaches that use a mechanism to deal with concept drift [1, 6, 7, 18, 19].
- Approaches that do not explicitly use a mechanism to detect drifts [4, 10, 11, 24].

Former category of online approaches use some measure related to the accuracy to handle drifts. They rebuild the system once a drift is detected/confirmed, and so cannot handle recurrent or predictable drift. These approaches suffer from non-accurate drift detections but respond quickly to changing concepts.

The latter set of approaches assigns weights to each base learner according to its accuracy, allows deletion of poor performing classifiers and add newly learnt classifiers. These approaches take longer time to recover from drifts but give results that takes into account the earlier learning and experience.

An ensemble of experts is a set of experts whose individual decisions are combined by weighted or un-weighted voting or by maximum value to classify new examples. The success of an ensemble in depicting the label of the new training example depends on the diversity of the base-experts.

Table 1. Yule's q-statistic classification table to measure diversity among the experts.

Number of examples	Classification by D_i	Classification by D_k
N^{11}	1	1
N^{10}	1	0
N^{00}	0	0
N^{01}	0	1

"Diversity "is the measure of variation in the classification accuracy of ensemble members for a given training example. In case of un-stable concepts, there exists a positive correlation between accuracy of the ensemble and diversity among its members [15, 22] but a negative correlation exists between the two for stable concepts. Considering two classifiers D_i and D_k, the main diversity measure used is Yule's Q-Statistic, given in Equation 1 and the corresponding classification is given as in Table 1.

$$Q_{i,k} = (N^{11} N^{00} - N^{01} N^{10}) / (N^{11} N^{00} + N^{01} N^{10}) \quad (1)$$

Higher value of Q-average means lower diversity and lower value of Q-average means higher value of diversity. Different levels of diversity in an ensemble were ensured by varying the value of λ in a modified version [15] of online bagging [21], where Poisson (1) distribution in Online Bagging has been replaced by

Poisson (λ) distribution.

Higher/lower values of λ are associated with lower /higher diversity in an ensemble of experts. The term prequential [4] accuracy defines the average accuracy obtained by the prediction of each example to be learned, before its learning, calculated in an online way.

In our paper, we will be developing a new online approach to handle drifting concepts in data streams based on diversity and the ensembles of weighted experts. In the section 2, we discuss the various online ensemble approaches to handle drifting concepts. In section 3, a study of the datasets to be used for empirical evaluation would be done. In section 4, a thorough study of our proposed approach would be conducted followed by empirical evaluation of our approach using various datasets in section 5. In the end, we summarize our paper and discuss directions for future research.

2. Related Work

Weighted Majority (WM) [3, 13] is an online approach which predicts using a set of weighted experts. It states that not all features are necessary to make a final prediction. Based on the maximum of weights of experts that predicted an incorrect output and experts that predicted a correct output, the algorithm uses WM voting to make a final prediction.

DWM [10, 12] is a modified version of WM [13]. It dynamically creates and updates the experts in response to changes in its global performance, and removes an expert if its weight reaches a threshold value.

Drift Detection Method (DDM) [6], controls the online error-rate (number of errors) produced during prediction. DDM adopts a dynamic window structure which is reduced when the error rate increases and is increased when there is a reduction in error rate. To overcome the limitation of DDM in handling very slow gradual drifts, Early Drift Detection Method (EDDM) [1] was proposed. It was based on the calculated distribution of the distances between classification errors.

Adaptive Classifier Ensemble (ACE) [20] was a classifier ensemble that uses an online classifier, a set of batch classifiers, and a drift detection mechanism to handle mainly recurrent drifts. ACE and DWM respond to sudden changes very quickly. An enhanced version of ACE [18] was introduced, with an improved weighting method. Experiments have shown that it responds to sudden changes more quickly and more accurately than the original version and the added pruning method helped it to retain the useful classifiers.

Detection with Statistical Test of Equal Proportions (STEPD) [19] is a single online classifier system that compares the overall accuracy from the beginning of the learning with the accuracy of recent examples after concept drift, by using statistical test of equal proportions.

Addictive Expert ensembles (AddExp) [11] adds a new online classifier whenever the system output is incorrect, and its weight is the total weight of the ensemble times a constant $\gamma \in (0, 1)$. Two pruning methods were proposed: the oldest first and the weakest first.

Two online classifiers for learning and detecting concept drift (Todi) [17] was developed to reduce the impact of false alarms on the average prequential accuracy. It uses two online classifiers (H_0 and H_1) for handling drifts. After the drift is detected, one of the classifiers is reinitialized (H_0) while the other one (H_1) is not.

Diversity for Dealing with Drifts (DDD) [16] used the concept of varying diversity levels between ensembles for the first time. Before the detection of drift there were two ensembles: low and high diversity ensembles which were both used for training but only the low diversity ensemble was used for making predictions. After the detection of drift, new high and low diversity ensembles were created and the earlier ensembles were denominated as old low and old high diversity ensembles. The system predictions were the WM vote of the output of the three ensembles: old low, old high and new low diversity ensemble. In case of stable concepts, DDD has almost similar accuracy as that of DWM and EDDM, but attains higher accuracy as compared to EDDM and DWM for continuous drifts.

3. Concept Drifting Data Sets

Empirical analysis of our approach was done using SEA dataset, hyperplane dataset and electricity pricing domain to the other concept drifting approaches such as WM [13] and DWM [10] in MOA [2], a tool for analyzing online data streams.

3.1. Moving Hyperplane Dataset

A hyperplane [9], in a d-dimensional space is a set of points **s**, that satisfy the condition as in Equation 2.

$$\sum_{i=1}^{d} a_i s_i = a_0, \tag{2}$$

Where $s_i \in [0, 1]$ is the i^{th} coordinate of **s** and a_i, represents the weights in each dimension i. Classification is positive as per the condition defined in Equation 3 else it is negative.

$$\sum_{i=1}^{d} a_i s_i \geq a_0, \tag{3}$$

The hyperplanes are useful for simulating gradually drifting concepts because we can easily change their orientation and position by changing their relative weights. Gradual changes were introduced by reversing the sign of inequality as in Equation 3, after N examples in dataset. Noise was introduced by switching the labels of 5% of the training examples.

3.2. SEA Concepts

The SEA concepts [25], provides a very large dataset containing sudden drift. Each example consists of three real-valued attributes, $x_i \in \mathbb{R}$ such that $0.0 \leq x_i \leq 10.0$.The target concept is as in Equation 4, for each of the four data blocks.

$$y = x_0 + x_1, \tag{4}$$

Each example belongs to class 1 if the condition in Equation 5 is true else it belongs to class 0.

$$x_0 + x_1 \leq \theta, \tag{5}$$

Where $\theta \in \{7, 8, 9, 9.5\}$ (one for each of the four data blocks). Thus, only the first two attributes of every instance are relevant and the third attribute is irrelevant. The SEA dataset was induced with 10% class noise for all the experimental evaluations.

3.3. Electricity Pricing Domain

To evaluate our system on a real world problem, we selected the electricity pricing domain [8]. Harries obtained this dataset from TransGrid, the electricity supplier in New South Wales, Australia. Every instance consists of five attributes and has a class label of either up or down. The attributes “day of week” and “period of day” have integer value in [1, 7] and [1, 48], respectively. The other three attributes measuring the current demand are: the demand in New South Wales, the demand in Victoria and the amount of electricity scheduled for transfer between the two states are numeric. The prediction task is to predict whether the price of electricity will go up or down based on the all these five attribute values.

For comparative analysis using hyperplane dataset, we generate a training set and a test set consisting of total of 10,000 examples. In case of SEA concepts, being a very large dataset we generate 50000 examples in totality. The electricity pricing dataset consists of 45,312 instances collected at 30-minute intervals from 7 May, 1996 to 5 December, 1998. We randomly generate one-fourth of the examples for testing and the remaining examples are used for training. At each time step, an online learning system was presented with one example, tested the concept descriptions using examples in the test set and computed the percent correctly predicted. In our experiments, the accuracy was calculated averaged over 40 trials.

4. Proposed Approach: Extended Dynamic Weighted Majority (EDWM)

In our approach, before drift detection two ensembles were created using different values of λ. One of the ensembles has higher diversity and the other ensemble has lower diversity created by using lower and higher values of λ, respectively in Poisson (λ) distribution of modified Online Bagging. We decide to use a DDM, DWM to treat drifts immediately once they are detected. Each of the experts in an ensemble is treated as weighted experts with an initial positive weight of one.

Both the ensembles are trained with each incoming example but only low diversity ensemble is used for prediction, as a prediction by a high diversity ensemble before drift detection may lead to reduced generalization accuracy. When a drift was detected, each of the ensembles are updated as per the DWM [10] approach. The weights of experts in the ensembles are reduced by a factor β if they gave an incorrect local prediction, after every p instances. An expert is removed from the ensemble if its weight reduces below a given threshold value, θ. A new expert is created in the ensembles when the WM vote prediction i.e., global prediction results are incorrect.

After drift was detected, EDWM allows the use of high diversity ensemble in form of an old high diversity ensemble. Two new ensembles, new low diversity and new high diversity ensembles are created as per training on the new concept. The ensembles before drift detection are kept and treated as old high and old low diversity ensembles. The old high diversity ensemble starts to learn with low diversity to easily adapt to the new concept. Both the old ensembles and new ensembles perform learning but the prediction is the WM global output of the old high, old low and new low diversity ensemble.

Our paper is an extension to our earlier work as in [23]. Our approach uses a DDM that does not explicitly uses a mechanism to deal with concept drift so, the need to decide the measure for accuracy is no longer there. It automatically adjusts the ensemble size in contrast to earlier approaches (DDD), where the size of ensemble had to be explicitly defined. Our approach can obtain good accuracy if the recurrent drifts are very frequent so that the weights of the base learners do not decay enough for them to be eliminated. Next, our approach automatically prunes or adds new classifiers according to their accuracy on new data instances, so it would not suffer from non-accurate drift detections in contrast to the earlier approaches. Next we present the empirical evaluation of our approach in contrast to WM and DWM using the above mentioned datasets.

5. Experimental Study and Results

We evaluated EDWM with naive bayes (EDWM-NB), DWM with naive bayes (DWM-NB) and WM using naïve bayes (WM-NB) with pairs of features as experts with the value of β=0.5. The value of threshold θ was considered to be 0.01 and ensemble size to be 10. For WM-NB, the value of k was set to one, which would make it react faster to concept drift using pair of features as experts.

5.1. Performance on Moving Hyperplane Dataset

For empirical evaluation on moving hyperplane dataset, the value of the parameter p that controlled the expert creation and weight update is set to 50. After lot of experimentation, the most effective value of λ for low diversity ensemble is one and for high diversity ensemble the chosen value was 0.0001.

Table 2. Experimental results for EDWM-NB, WM-NB and DWM-NB on moving hyperplane dataset averaged over 40 trials.

	EDWM-NB	DWM-NB	WM-NB
Accuracy (%)	85.86	81.38	85.81
kappa statistic (%)	71.50	67.05	71.40
RAM-Hours	0.00	0.00	0.00
Time (CPU-seconds)	5.16	3.18	10.7
Memory(bytes)	0.03	0.01	0.02

Our approach, EDWM-NB gives the best accuracy among all the three systems as seen in Table 2. The generalization accuracy of EDWM and WM is almost about 86% whereas DWM depicts the least accuracy. Our system gives better average prequential accuracy than DWM at all time steps as shown in Figure 1.

EDWM is more robust to noise in contrast to DWM as observed by large number of variations in accuracy in the graph of DWM.

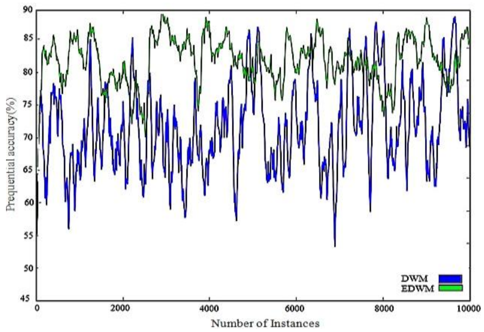

Figure 1. Prequential accuracy curves for EDWM and DWM on moving hyperplane dataset.

The Kappa statistic value is a performance measure that gives a score of homogeneity among the experts. To get the optimum results for drifting concept distributions we need the experts to be diverse from each other such that if one expert classifies an instance incorrectly, another expert gives a correct classification. As per our observation of results in Table 2, the experts of DWM are more diverse than our approach but as in-case of gradual drifts, high diversity reduces accuracy, so EDWM gives better accuracy than DWM with optimum level of diversity.

The resource efficiency of any stream mining algorithm, is judged by another new measure i.e., RAM-Hours. One RAM-Hour is equivalent to one GB of RAM being deployed for one hour. Hence the analysis of the results, conclude that all the three online systems are highly resource-efficient making them extremely suitable for any real time application.

The total CPU involvement in updating the ensembles is higher for our system than DWM as it handles four kinds of ensembles and involves more processing to update and create/ delete the experts. Similarly, the memory needs of our system are higher than DWM as it maintains more number of experts at any given time step. WM takes more time in contrast to our approach as it maintains no limit on the number of experts and does not prune any poor performing experts. Hence, when implemented on dataset with gradual change, our approach gives very good accuracy within real time.

5.2. Performance on SEA Concepts

The value of the parameter that controlled the expert creation and weight update i.e., p=4. To get the best results, the most effective value of λ for lower diversity ensemble was 1 and for higher diversity ensemble was 0.001.

The experimental results for EDWM, DWM and WM on very large dataset, i.e., SEA concepts averaged over 40 trials have been provided in Table 3.

Table 3. Experimental results for EDWM-NB, WM-NB and DWM-NB on SEA Concepts averaged over 40 trials.

	EDWM-NB	DWM-NB	WM-NB
Accuracy (%)	85.43	83.66	83.23
kappa statistic (%)	66.16	65.14	67.71
RAM-Hours	0.00	0.00	0.00
Time (CPU-seconds)	80.70	87.34	74.60
Memory(bytes)	0.02	0.01	0.02

Our approach, EDWM best generalizes SEA concepts as compared to DWM and WM. This is because, if the lower diversity ensemble does not predict correctly we have a very good probability of correct prediction by the higher diversity ensemble. Hence, if one ensemble does not give good accuracy, the other ensemble accommodates that and gives good accuracy levels. EDWM reacts earlier to drifts in concept as compared to DWM which detects drifts later as illustrated in Figure 2, just before time steps 12500 and 25000. Our approach reaches target concepts earlier than DWM approach and that too with higher accuracy as illustrated in the fourth target concept just after time steps 37500.

The experts in EDWM gave good diversity levels almost similar to the other weighted approaches as stated by the value of kappa statistic. Again, the value of RAM-Hours supports the fact that all the systems in any kind of domain are highly resource efficient, suitable for a real time domain with drifting concepts.

In terms of total evaluation time performance parameter, our approach scores better than DWM approach. This is because the frequency of classification errors is higher in-case of DWM in contrast to our approach and hence more CPU involvement to update, create or delete the poor performing experts.

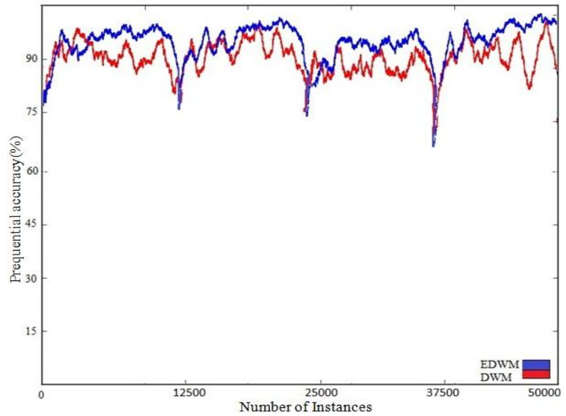

Figure 2. Prequential accuracy curves for EDWM and DWM on SEA, very large dataset with noise.

EDWM and WM perform similarly in terms of memory usage as they have to store large number of experts. On the other hand, the memory needs of DWM are almost half as it prunes poor performing experts whose weight is below a defined threshold value and creates new experts only when the global prediction results are incorrect.

Empirical evaluation of EDWM on two artificial data sets: one with gradual drift and another one with sudden drifts containing noise clearly prove that our approach achieves very good accuracy in contrast to the other online weighted approaches with a slight increased time and memory requirements. EDWM is highly robust to noise and maintains ensembles with an optimum level of diversity.

5.3. Performance on Electricity Pricing Domain

For empirical evaluation the value of the parameter *p* that controlled the expert creation and weight update is set to 10. The most effective value of λ for lower diversity ensemble was found to be 1 and for higher diversity ensemble was 0.01. Specifically for this dataset, the ensemble size is taken to be 5.

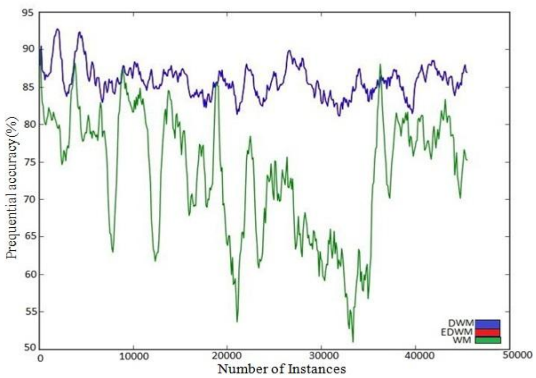

Figure 3. Prequential accuracy curves for EDWM, DWM and WM on electricity pricing domain.

EDWM and DWM show similar prequential accuracy as seen by the overlapping of their graphs as seen in Figure 3. Both these systems reach target concepts at the same time and with same frequency. EDWM however takes more time as compared to DWM as seen in Figure 4. Both these systems depict linear time graphs.

Our approach, EDWM shows better accuracy than the WM approach as seen in Figure 3. EDWM shows almost consistent accuracy and the system is highly robust to change as compared to WM approach that depicts higher variations in accuracy. Our approach however takes more time as compared to WM as seen in Figure 4 but using lesser space than the WM approach. The experimental results averaged over 40 trials for all the approaches on Electricity pricing domain have been provided in Table 4.

This dataset being derived from a real world scenario, we are not clear if and when drifts occur and what type of drift occurs. However, experimental results have proved that EDWM provides us with a resource efficient concept drifting system better than the WM approach and almost similar as DWM approach.

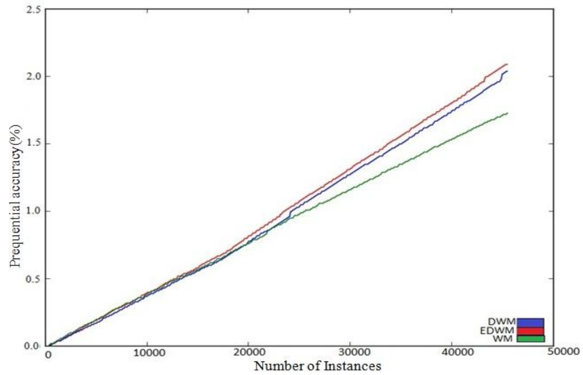

Figure 4. Evaluation time graphs for EDWM, DWM and WM on Electricity pricing domain.

Table 4. Experimental results for EDWM-NB, WM-NB and DWM-NB on Electricity pricing domain averaged over 40 trials.

	EDWM-NB	DWM-NB	WM-NB
Accuracy (%)	85.86	85.86	73.51
kappa statistic (%)	70.68	70.68	39.93
RAM-Hours	0.00	0.00	0.00
Time (CPU-seconds)	0.99	0.95	0.88
Memory(bytes)	0.01	0.01	0.03

6. Summary and Conclusions

Tracking concept drift is important nowadays to predict the future trends. The study of drifting concepts and continuous learning has helped to predict the behaviour of various real time data streams. Our approach, EDWM handled gradual as well as sudden drifts with very good accuracy levels within real time and memory. It proved to be very resource effective, which is the need of the hour for any real time domain.

For future work, we can extend it to handle recurrent drifts, where lot of scope for research is still open. We could also use other diversity measures to calculate the diversity among the experts. Different ways to assign and update weights of the experts should be developed. We should take measures to improve our system and make it more space and time efficient. Further, studies could be conducted using different base learning algorithms other than naïve bayes classifier.

References

[1] Baena-Garcı´a M., Campo-Avila J., Fidalgo R., Bifet A., "Early Drift Detection Method," *in Proceeding of 4th International Workshop on Knowledge Discovery from Data Streams*, pp. 77-86, 2006.

[2] Bifet A., Holmes G., Kirkby R., and Pfahringer B., "MOA: Massive Online Analysis, a Framework for Stream Classification and Clustering," *in Proceeding of Workshop on Applications of Pattern Analysis*, Windsor, pp. 44-51, 2010.

[3] Blum A., "Empirical Support for Winnow And Weighted-Majority Algorithms: Results on a Calendar Scheduling Domain," *Machine Learning*, vol. 26, no. 1, pp. 5-23, 1997.

[4] Dawid A. and Vovk V., "Prequential Probability: Principles and Proper Ties," *Bernoulli*, vol. 5, no. 1, pp. 125-162, 1999.

[5] Dietterich T., "Machine Learning Research: Four Current Directions," *Artificial Intelligence*, vol. 18, no. 4, pp. 97-136, 1997.

[6] Gama J., Medas P., Castillo G., and Rodrigues P., "Learning with Drift Detection," *in Proceeding Brazilian Symposium on Artificial Intelligence*, Sao Luis, pp. 286-295, 2004.

[7] Gao J., Fan W., and Han J., "On Appropriate Assumptions to Mine Data Streams: Analysis and Practice," *in Proceeding of IEEE International Conference on Data Mining*, Omaha, pp. 143-152, 2007.

[8] Harries M., Splice-2 Comparative Evaluation: Electricity Pricing, Technical Report, University of New South Wales, 1999.

[9] Hulten G., Spencer L., and Domingos P., "Mining Time-Changing Data Streams," *in Proceeding of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, pp. 97-106, 2001.

[10] Kolter J. and Maloof M., "Dynamic Weighted Majority: A New Ensemble Method for Tracking Concept Drift," *in Proceeding of the 3rd IEEE International Conference on Data Mining*, Melbourne, pp. 123-130, 2003.

[11] Kolter J. and Maloof M., "Using Additive Expert Ensembles To Cope With Concept Drift," *in Proceeding of the Twenty Second ACM International Conference on Machine Learning*, Bonn, pp. 449-456, 2005.

[12] Kolter J. and Maloof M., "Dynamic Weighted Majority: An Ensemble Method for Drifting Concepts," *The Journal of Machine Learning Research*, vol. 8, pp. 2755-2790, 2007.

[13] Littlestone N. and Warmuth M., "The Weighted Majority Algorithm," *Information and Computation*, vol. 108, no. 2, pp. 212-261, 1994.

[14] Mansoori M., Zakaria O., and Gani A., "Improving Exposure of Intrusion Deception System through Implementation of Hybrid Honeypot," *The International Arab Journal of Information Technology*, vol. 9, no. 5, pp. 436-444, 2012.

[15] Minku L., White A., and Yao X., "The Impact of Diversity on On-Line Ensemble Learning in the Presence of Concept Drift," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 5, pp. 730-742, 2010.

[16] Minku L. and Yao X., "DDD: A New Ensemble Approach for Dealing with Concept Drift," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 4, pp. 619-633, 2012.

[17] Nishida K., "Learning and Detecting Concept Drift," *PhD Dissertation*, Hokkaido University, 2008.

[18] Nishida K. and Yamauchi K., "Adaptive Classifiers-Ensemble System For Tracking Concept Drift," *in Proceeding of the 6th International Conference on Machine Learning and Cybernetics*, Honk Kong, pp. 3607-3612, 2007.

[19] Nishida K. and Yamauchi K., "Detecting Concept Drift Using Statistical Testing," *in Proceeding of the 10th International Conference on Discovery Science*, Sendai, pp. 264-269, 2007.

[20] Nishida K., Yamauchi K., and Omori T., "ACE: Adaptive Classifiers-Ensemble System for Concept-Drifting Environments," *in Proceeding of the 6th International International Conference on Multiple Classifier Systems*, California, pp. 176-185, 2005.

[21] Oza N. and Russell S., "Experimental Comparisons of Online and Batch Versions of Bagging and Boosting," *in Proceeding of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, pp. 359-364, 2001.

[22] Scholz M. and Klinkenberg R., "An Ensemble Classifier for Drifting Concepts," *in Proceeding of the Second International Workshop on Knowledge Discovery from Data Streams*, Porto, pp. 53-64, 2005.

[23] Sidhu P. and Bhatia M., "Extended Dynamic Weighted Majority Using Diversity to Handle Drifts," *in Proceeding of 17th East European*

Conference on Advances in Databases and Information Systems, pp. 389-395, 2014.

[24] Stanley K., “Learning Concept Drift with A Committee of Decision Trees,” Technical Report UT-AI-TR-03-302, Department of Computer Sciences, University of Texas at Austin, 2003.

[25] Street W. and Kim Y., “A Streaming Ensemble Algorithm (SEA) for Large-Scale Classification,” *in Proceeding of 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, pp. 377-382. NY, 2001.

Parneeta Sidhu received her B.Tech degree in Computer Science from Punjab Technical University in 2002. She received her M. Tech degree in Information Systems from University of Delhi in 2009. Ms. Sidhu is a Teaching cum Research Faculty in the Division of COE at the Netaji Subhas Institute of Technology, affiliated to University of Delhi. She is presently pursuing her PhD in Computer Science from University of Delhi under the guidance of Dr. MPS Bhatia. Her research interests include data mining, concept drift, outlier analysis in data streams. She is an author or coauthor of 8 research papers in various international journals and conferences of high repute. She is a member of CSI (Computer Society of India).

Mohinder Bhatia received his PhD in Computer Science from University of Delhi. Dr. Bhatia is a Professor in the Division of COE at the Netaji Subhas Institute of Technology, affiliated to University of Delhi. He is also serving the Institute as Dean, Student Welfare and Head, and Head, Placement Cell. He has guided many M.Tech and PhD students in their research work. His research interests include data mining, cyber security, semantic web, machine learning, social network analysis and sentiment analysis. He is an author or coauthor of many research papers in international journals and conferences. Dr. Bhatia is a member of IEEE (Institute of Electrical and Electronics Engineers) and CSI (Computer Society of India).