

# Ensemble based on Accuracy and Diversity Weighting for Evolving Data Streams

Yange Sun<sup>1</sup>, Han Shao<sup>1</sup>, and Bencai Zhang<sup>2</sup>

<sup>1</sup>School of Computer and Information Technology, Xinyang Normal University, China

<sup>2</sup>School of Computer and Information Technology, Beijing Jiaotong University, China

**Abstract:** Ensemble classification is an actively researched paradigm that has received much attention due to increasing real-world applications. The crucial issue of ensemble learning is to construct a pool of base classifiers with accuracy and diversity. In this paper, unlike conventional data-streams oriented ensemble methods, we propose a novel Measure via both Accuracy and Diversity (MAD) instead of one of them to supervise ensemble learning. Based on MAD, a novel online ensemble method called Accuracy and Diversity weighted Ensemble (ADE) effectively handles concept drift in data streams. ADE mainly uses the following three steps to construct a concept-drift oriented ensemble: for the current data window, 1) a new base classifier is constructed based on the current concept when drift detect, 2) MAD is used to measure the performance of ensemble members, and 3) a newly built classifier replaces the worst base classifier. If the newly constructed classifier is the worst one, the replacement has not occurred. Comparing with the state-of-art algorithms, ADE exceeds the current best-related algorithm by 2.38% in average classification accuracy. Experimental results show that the proposed method can effectively adapt to different types of drifts.

**Keywords:** Ensemble learning, concept drift, accuracy, diversity.

Received March 27, 2020; accepted February 21, 2021  
<https://doi.org/10.34028/iajit/19/1/11>

## 1. Introduction

Data stream classification has gained growing attention owing to its wide applications. Examples include credit card fraud detection, spam filtering, intrusion detection, and data analysis in the Internet of Things (IoT) networks [1, 8, 9, 12, 17]. Building a classifier incrementally under data stream scenarios is one of the most critical issues in data mining. However, data stream is usually nonstationary. That is, the distribution of data may evolve dynamically over time. This phenomenon leads to the issue of concept drift [7, 15]. The occurrence of concept drift often degrades the performance of the classification model, which brings challenges to the incremental learning process of data streams.

Adaptation to concept drift is also a vital issue for the study of unstable data streams. Many techniques have been proposed to deal with the challenge [7, 10, 13, 21, 24, 25, 26]. Among these efforts, ensemble approaches [10] are one of the most frequently used methods due to their flexibility and ability to enhance the performance of learning algorithms. Ensemble methods are updated based on newly arrived data, thus providing a natural mechanism to update the knowledge base by adding new classifiers (concepts), moving, or updating outdated classifiers to adapt to the evolving environment [10].

Most existing ensemble methods for data stream adapt to concept drift by constantly updating ensemble according to the performance of the base classifier. The

critical to a successful ensemble is to create ensemble members with high performance and high diversity [3, 5, 14, 22]. In addition, the issue of concept drift requires ensembles to consider both accuracy and diversity during the construction process simultaneously. It will make the algorithm stable and adaptive. That is, it can quickly recover and adapt to new concepts after a drift occurs. However, most existing ensemble-based data stream classification methods only employ accuracy or diversity to evaluate the performance of the base classifier in the ensemble construction process. Therefore, it is meaningful to construct a more effective ensemble classification by injecting two metrics simultaneously.

Combining two metrics to propose a more effective ensemble schema is the main task of this paper. Therefore, a novel ensemble approach, called Accuracy and Diversity weighted Ensemble (ADE), is devised to handling various concept drifts. The contributions of this paper can be summarized as follows:

1. We tailor the ensemble learning to handling different kinds of concept drifts by equipping with explicit change detection mechanisms and dynamic weighting schema.
2. A dynamic weighting mechanism based on accuracy and diversity is proposed to boost the generalization performance of ensemble.
3. The performance of our algorithm was implemented on different kinds of data stream

datasets. The results demonstrated that ADE achieves the best overall performance and drift adaption.

The paper is organized as follows. Section 2 retrospects related work. Section 3 explains the basic idea of our approach and presents its details. The algorithm is analyzed and experimentally evaluated on real and synthetic datasets in section 4. A conclusion is drawn in section 5.

## 2. Related Work

For most data stream ensemble methods, base classifiers are weighted and selected according to the accuracy. Accuracy Weighted Ensemble (AWE) [25] is the most famous representative ensemble based on accuracy. In AWE, base classifiers are trained in sequence from data blocks, and base classifiers are dynamically weighted according to the accuracy of the newly arrived block. Brzezinski et al. proposed an ensemble based on accuracy called Accurate Updated Ensemble2 (AUE2) [4]. Pietruczuk *et al.* [18] proposed an ensemble method called Dynamically Expanded Ensemble Algorithm (DEEA), which adapts the size of the ensemble according to the overall accuracy of the ensemble.

In addition to the measure of accuracy, diversity is another meaningful measure for data stream classification. Diversity is a critical factor in improving ensemble performance [3, 22]. A set of bagging ensembles based on diversity, named Diversity for Dealing with Drifts (DDD), is presented by Minku and Yao [16]. They argue that taking advantage of diversity is beneficial to deal with concept drift. Santos *et al.* [20] proposed an online boosting method based on diversity. Adaptive Random Forest (ARF) [11] is an improved random forest algorithm based on diversity. Recently, Sun *et al.* [23] presented the Diversity and Transfer-based Ensemble Learning (DTEL) algorithm, which utilized  $Q$  statistic to measure diversity to determine whether the previous base classifier is retained. Rijn *et al.* [19] proposed a heterogeneous ensemble model for data stream classification and weighting the voting of different base classifiers in the ensemble classifier. Kappa Update Ensemble (KUE) utilized Kappa statistics to dynamically update base classifiers' weights [6].

The generalization of ensembles largely depends on both accuracy and diversity. Unfortunately, most of the above ensembles are based only on one factor. Our approach is different from the above methods in the following aspects: First, we adopt a dynamic ensemble structure, using the selective ensemble strategy to react both abrupt and gradual drifts simultaneously. Second, we devise an efficient ensemble that considers accuracy and diversity.

## 3. Our Approach

### 3.1. Notations and Basic Idea

Let  $S$  denotes a stream,  $W=\{B_i|i=1, 2, \dots, M\}$  denotes data stream with each data block  $B_i=\{(x_j, y_j)|j=1, 2, \dots, N\}$ , where  $y_j \in \{1, 2, \dots, K\}$  is the class of instance  $x_j$  and  $|B_i|=|B_k|$  for  $i \neq k$ .

The proposed method aims to construct a dynamic ensemble or a set of classifiers  $E=\{C_i|i=1, 2, \dots, n\}$  from sequential blocks of data stream  $S$ . More specifically, whenever a new block arrives, or change detects, we build a new classifier from it. Then the classifiers are dynamic weighted based on their classification accuracy and diversity in a changing environment. Top- $k$  classifiers are used to construct the ensemble. The final decision of class label is based on weighted majority votes of base classifiers.

### 3.2. Weighting based on Accuracy and Diversity

The performance of ensembles relies on member classifiers' accuracy and diversity. However, few approaches inject accuracy and diversity into an ensemble. To consider both the accuracy and diversity of a classifier, we adopt a novel weighting strategy that calculates the weights of classifiers according to both accuracy and diversity.

When a new data block arrives, a new classifier  $C'$  is built based on it. The weight of  $C'$  ( $w_{C'}$ ) is calculated according to Equation (1).

$$w_{C'} = \alpha * \text{Acc}_{C'} + (1 - \alpha) * \text{div}_{C'} \quad (1)$$

Where  $\alpha \in [0, 1]$  representing the weight of accuracy.

The weight of the base classifier also needs to adjust according to the new data block and using Equation (2) to calculate the new weight value of the base classifier.

$$w_{ij} = \alpha * \text{Acc}_{ij} + (1 - \alpha) * \text{div}_{ij} \quad (2)$$

Where  $w_{ij}$  represents the weight of base classifier  $C_i$  on data block  $B_j$ .

1) The measure of accuracy

The measure of accuracy is calculated by the Mean Squared Error (MSE). The mean square error  $MSE_{ij}$  of a classifier  $C_i$  in a block  $B_j$ , can be calculated according to Equation (3).

$$MSE_{ij} = \frac{1}{|B_j|} \sum_{\{\mathbf{x}, y\} \in B_j} (1 - p(f_i(\mathbf{x}) = y))^2 \quad (3)$$

Let  $MSE_r$  represents the mean squared error of random predictions to reflect the distribution of current data.  $MSE_r$  is calculated according to Equation (4).

$$MSE_r = \sum_y p(y)(1 - p(y))^2 \quad (4)$$

The accuracy can be expressed in  $MSE_{ij}$  and  $MSE_r$ . This paper uses Equation (5) to represent the base classifiers' accuracy in the ensemble.

$$Acc_{ij} = \frac{1}{MSE_{ij} + MSE_r + \delta} \quad (5)$$

The existence of  $\delta$  is to prevent the denominator of Equation (5) from being zero. In addition, the accuracy of the new classifier  $Acc_{C'}$  is calculated according to Equation (6).

$$Acc_{ij} = \frac{1}{MSE_{ij} + MSE_r + \delta} \quad (6)$$

The existing theoretical and experimental studies show that ensemble classifiers are more generalized than single classifiers. Moreover, the ensemble of independent, complementary, and relatively accurate classifiers is superior to the best base classifier in the generalization performance [16]. When concept drift occurs, base classifiers may not be able to adapt to new concepts. If the base classifiers in the ensemble classifier are diverse, one can always find the best way to handle the new problem. So, the ensemble classifier has a good generalization.

## 2) The measure of diversity

In our algorithm, we adopt the  $Q$  statistic as the diversity measure.  $Q$  statistic value between two classifiers is calculated according to Equation (7):

$$Q_{ij} = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}} \quad (7)$$

Where  $N^{ab}$  represents the number of instances where  $C_i$  has a classification result  $a$  and  $C_j$  has a classification result  $b$ , 1 means the classification is correct, and 0 means the classification is wrong. For example,  $N^{11}$  indicates that the classifier  $i$  and  $j$  simultaneously classify the correct number of instances.

Then, we convert diversity into larger representational diversity, as shown in Equation (8).

$$Q_{ij}^* = 0.5 * (1 - Q_{ij}) \quad (8)$$

Where  $Q_{ij}$  represents the  $Q$  statistics of  $C_i$  and  $C_j$ , and uses the value to express the difference between the two classifiers. Since the range of  $Q_{ij}$  is  $[-1, 1]$ , the range of  $1 - Q_{ij}$  is  $[0, 2]$ . Then perform normalization, which has little effect on the final result.

The diversity value of one classifier  $C$  and ensemble  $E$  is an average value of  $Q_{ij}^*$  calculated by Equation (8) for  $C$  and each base classifier in  $E$ . Therefore,  $div_C$  denotes diversity between a newly established classifier  $C'$  and ensemble  $E$ .  $div_C$  can be calculated according to Equation (9).

$$div_{C'} = \sum_{i=C', C_j \in E} Q_{ij}^* / |E| \quad (9)$$

The diversity value of the set formed by one of the base classifiers  $C_i$ , and other base classifiers in the ensemble can be calculated according to Equation (10).

$$div_{ij} = \sum_{C_i \in E, C_j \in E, i \neq j} Q_{ij}^* / |E| \quad (10)$$

## 3.3. Ensemble Algorithm Based on Accuracy and Diversity

In order to deal with different kinds of concept drift, the ADE was introduced to achieve better generalization performance under concept drift environments. The main benefit of this approach is that the diversity of the final solution can be accurately measured and can take the trade-off between accuracy and diversity of the whole ensemble into account during ensemble generation.

For each new incoming data block  $B_i$ , a new classifier  $C'$  is first constructed based on  $B_i$  when change detect. Then the weight of  $C'$  is calculated according to Equation (1). Use Equation (2) to calculate the weight of each base classifier. The weights of the classifiers are obtained by linearly weighting their accuracy on the  $B_i$  of the latest data block and their diversity values in the ensemble. When ensemble is full, the base classifier with the smallest weight is replaced by a newly built classifier. ADE is listed as follows.

*Algorithm 1: Accuracy and Diversity weighted Ensemble*

*Input: Stream S, the maxsize of ensemble k*  
*Output: a pool of k weighted classifiers*  
 01: initialize E;  
 02: for every data block  $B_i \in S$  do  
 03: if  $|B_i| = d$  or change detected then  
 04: train a new classifier  $C'$  on data block  $B_i$ ;  
 05: compute the diversity and accuracy of  $C'$ ;  
 06: compute the weight of  $C'$  by the Equation (1);  
 07: for all classifiers  $C_j \in E$  do  
 08: compute the diversity and accuracy of  $C_j$ ;  
 09: compute the weight of  $C_j$  by the Equation (2);  
 10: if  $|E| < k$   
 11: add  $C'$  to the ensemble;  
 12: else  
 13: substitute the worst base classifier with  $C'$ ;  
 14: end if  
 15: end if  
 16: end for  
 17: for  $C_j \in E \setminus \{C'\}$  do  
 18: incrementally train  $C_j$  on  $B_i$ ;  
 19: end for  
 20: end for  
 21: end.

## 4. Experimental Evaluation

The experiments were implemented with Massive Online Analysis (MOA) [2] on a PC with 3.0 GHz CPU, 32 GB memory, and Windows 10. MOA provides the state-of-the-art data stream algorithms and tools for evaluation.

## 4.1. Datasets

In our experiments, we adopted eight benchmark datasets, including four synthetic and four real-world datasets. The streams were all generated by stream generators in the MOA framework to simulate different concept drift scenarios. The datasets are described in Table 1.

Table 1. Characteristics of the datasets.

Dataset	Instances	Attributes	Classes	Drift Type
HyperPlane	1,000,000	10	2	gradual
SEA	1,000,000	3	4	sudden
LED	1,000,000	24	10	mixed
Random Tree	1,000,000	10	4	sudden recurring
Electricity	45,312	10	10	-
Poker	1,000,000	10	10	-
Covertypes	581,012	53	7	-
Airlines	539,383	7	2	-

### 1) Synthetic data streams

The Hyperplane is used to simulate gradual concept drift. Change the decision boundary of the stream slowly by MOA generator to simulate gradual drift. In the Streaming Ensemble Algorithm (SEA) stream, data is divided into four blocks to represent different concepts. A sudden change from a class concept to another is utilized to simulate sudden concept drift. The Light Emitting Diode display (LED) is a stream with mixed drifts. The Random Tree stream contains sudden recurring drifts.

### 2) Real-world data streams

The Electricity dataset aims to predict the trend of price in the Australian New South Wales Electricity Market. The Poker dataset represents the problem of identifying the hand in a poker game. Each instance represents a hand comprising of five cards drawn from a standard deck of 52. The Covertypes dataset is to predict the forest cover types from cartographic variables. Airlines dataset contains information of predicting whether the flight will be delayed for given scheduled departure information. We employed the generators to transfer the static datasets into streams.

## 4.2. Results and Discussion

### 4.2.1. Parameter Sensitiveness

#### 1) Sensitivity analysis of the block size

Figure 1 illustrates the accuracy change of the proposed algorithm using different sizes of data blocks on Covertypes, SEA, and Random tree. The size of a block is an important factor in the performance of the block-based ensemble. Selecting a larger data block means using enough instances to train a base classifier, resulting in better classification performance. The disadvantage is that it is insensitive to the concept drift response. For example, the sudden change of the classifier can be greatly reduced by the sudden change

of the classifier. On the contrary, choosing a smaller block is more sensitive to drifts so that can better adjust the classification according to the current data distribution. However, the disadvantage is that there are fewer instances for learning a classifier each time, resulting in lower classification accuracy for each classifier. Therefore, defining an appropriate size of the block can boost the overall performance of ensemble.

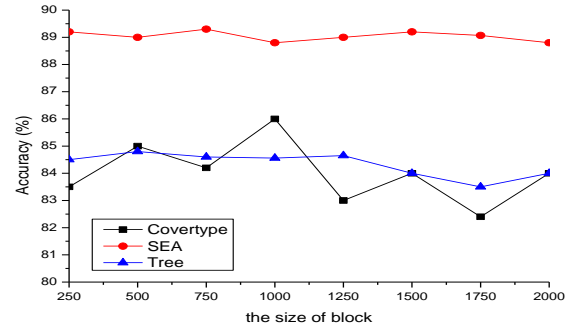


Figure 1. The effect of different data block sizes.

As shown in Figure 1, 1000 is the most appropriate block size for Covertypes, that is, to train a classifier using a data block for each 1000 instances arrival. The data block size of 500 is the most appropriate for SEA. We can see that different block sizes should be selected for different datasets to achieve the best performance.

#### 2) Sensitivity analysis for $\alpha$

The ADE calculates the weights of the classifiers by linearly weighting the diversity and precision. It can be seen from Equations (1) and (2) that controlling the proportion of diversity and accuracy in calculating the weights by  $\alpha$ , therefore, its value will have an impact on the final classification performance. Figure 2 shows different values of  $\alpha$  on the average classification accuracy on Poker and the LED datasets. We can see that different thresholds need to be selected to achieve the best classification performance for different datasets.

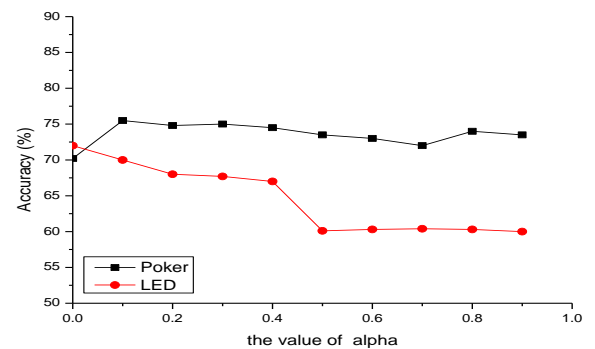


Figure 2. The effect of different  $\alpha$  on average accuracy.

### 4.2.2. Comparison Study

This section uses comparative experiments to verify the algorithm's effectiveness in terms of efficiency

and adaptability. ADE was evaluated against ARF, AUE2, and KUE. The average accuracy of the algorithms on different datasets is compared, as shown in Table 2.

Table 2. The accuracy of four algorithms (%).

	AUE2	ARF	KUE	ADE
<b>HyperPlane</b>	85.06 (4)	<b>91.5 (1)</b>	89.38 (2)	89.09 (3)
<b>Random Tree</b>	67.59 (3)	56.13 (4)	69.39 (2)	<b>69.95 (1)</b>
<b>SEA</b>	84.24 (3)	83.48 (4)	89.56 (1)	85.55 (2)
<b>LED</b>	76.74 (3)	76.22 (4)	78.47 (2)	<b>79.61 (1)</b>
<b>Coverttype</b>	70.21 (4)	86.50 (1)	80.27 (3)	86.29 (2)
<b>Poker</b>	68.28 (3)	67.53 (4)	69.67 (2)	<b>75.83 (1)</b>
<b>Airlines</b>	65.69 (4)	67.18 (2)	65.88 (3)	<b>72.25 (1)</b>
<b>Average</b>	73.97 (3.43)	75.51 (2.86)	77.52 (2.14)	<b>79.80 (1.57)</b>

1) Synthetic data streams

It can be seen from Table 2, ADE outperforms all the other algorithms in terms of accuracy, especially on the RandomTree, LED, Poker, and Airlines dataset. The ADE is the most accurate followed by KUE, and AUE2 achieves the worst performance. It is principally because ADE offers significant improvements over compare methods by dynamic accuracy and diversity weighting mechanism.

Figures 3-7 shows the curves of the classifications' accuracy with the increase of instances. Many kinds of literature adopt the graphical way to depict classification accuracy with changes in data distribution visually.

Figure 3 reveals the accuracy of each phase of the four algorithms incrementally trained on the SEA dataset. Since SEA is a dataset with abrupt drift, the data block size is set to 500 for responding concept drift quickly. The curve of the proposed method coincides with the curve of KUE, which shows that the two methods achieve similar performance in this scenario. Our method is only 0.31% higher than the KUE. This is partly because we tailored ensemble learning to handle concept drift by continuously updating the weights of base classifiers considering accuracy and diversity.

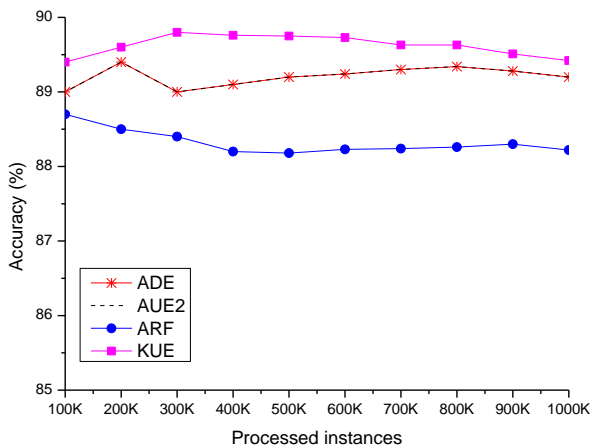


Figure 3. The accuracy on the SEA dataset.

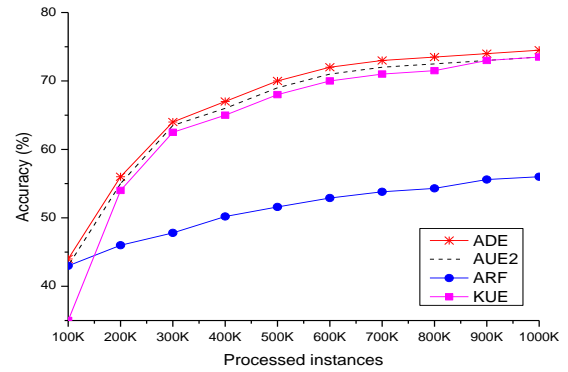


Figure 4. The accuracy on the RandomTree dataset.

Figure 4 illustrates the accuracy of the algorithms on the Random Tree stream. We generate the dataset with sudden drift by the RandomTreeGenerator in the MOA framework. As shown in Figure 4, ADE outperforms all the other algorithms, although it is slightly more accurate than AUE2. This is because the ADE dynamic weighting strategy uses the current and past classifiers combined with dynamically updated voting weights based on accuracy and diversity.

2) Real-world data streams

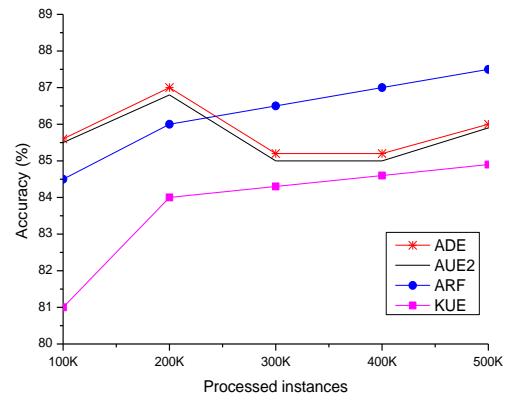


Figure 5. The accuracy on the Coverttype dataset.

Figure 5 shows the accuracy of the algorithms changes with the processed instances on the Coverttype scenario. ADE and AUE2 perform better in the early training stage, and ARF is better in the late training period. From the average accuracy, the difference between ADE was only 0.21% higher than the ARF. An analysis quickly reveals the cause: ADE benefits from accuracy and diversity to perform better in a real-world dynamic environment.

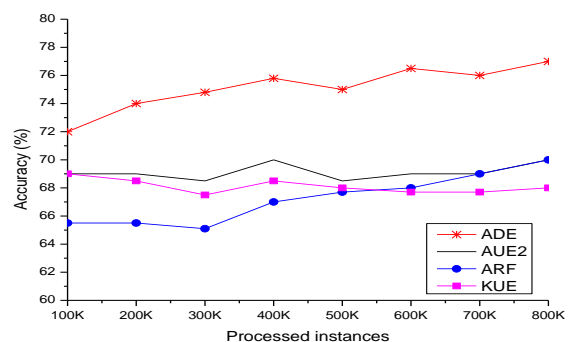


Figure 6. The accuracy on the poker dataset.

Figure 6 shows the accuracy of each phase of the four algorithms incremental training on the Poker dataset. It indicates that ADE is much better than the other three algorithms. This might be attributed to the weighting scheme based on accuracy and diversity can significantly boost the ensemble's performance. Hence, our algorithm achieves better performance in evolving real-world scenarios. It also verifies that the proposed ADE has a stronger ability to adapt to the real-world environment than other algorithms.

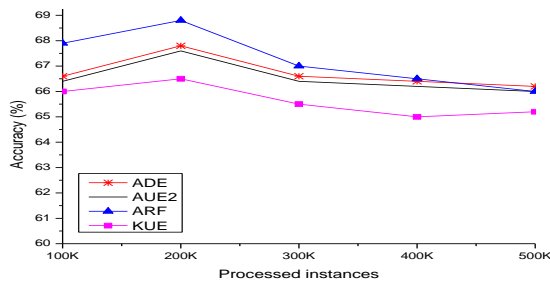


Figure 7. The accuracy on the airlines dataset.

Figure 7 indicates the accuracy of each of the four algorithms when incrementally training on the Airlines. Although the average precision of ARF from Table 2 is higher, it can be seen from Figure 7 that ARF has high average accuracy because it starts with high precision. As the instances continue to increase, the accuracy continues to decline. After 300,000 instances, ADE consistently achieves the best performance. This is partly because the weight of the classifier is updated based on the classification accuracy and classification error of the latest data, which can enhance the performance of the ensemble.

In conclusion, the experimental results show that ADE can achieve better average ranks than state-of-art in different drift scenarios, including abrupt, gradual, and recurring. It is mainly because ADE fully considers the two factors of accuracy and diversity when weighting and selecting the base classifiers. Hence, ADE achieves high performance and strong adaptability.

## 5. Conclusions

This paper focuses on the topic of constructing efficient ensemble under evolving data streams scenarios. Specifically, we design an effective metric based on accuracy and diversity to guide the construction of ensemble by evaluating the performance of base classifiers. Linearly weighting is employed to calculate the weights of classifiers. The higher weight indicates that the classifier is more important, and when the new classifier is established, it needs to replace the base classifier with the lowest weight. Experiments prove that the proposed algorithm can effectively adapt to scenes containing different kinds of concept drifts and quickly recover its performance.

Some limitations of this study are as follows: First, our algorithm's performance depends on the two parameters: the number of base classifiers  $k$  and the weight of accuracy  $\alpha$ . We will focus on automatically determine the parameters in the future. Second, recurring concept drifts are common in real-world applications, and we are devoted to considering this situation in our future work.

## Acknowledgments

This work is supported by National Natural Science Foundation of China (No. 62062004, 61702550, 31900710), Teacher Education Curriculum Reform projects of Henan Province (No. 2020-JSJYYB-034), the Science and Technology Project of Henan Province (No. 212102310993) and Nanhu Scholars Program for Young Scholars of XYNU.

## References

- [1] Aggarwal C., *Data Streams: Models and Algorithms*, Berlin: Springer-Verlag, 2007.
- [2] Bifet A., Holmes G., and Kirkby R., and Pfahringer B., "MOA: Massive Online Analysis," *Journal of Machine Learning Research*, vol. 11, no. 52, pp. 1601-1604, 2010.
- [3] Brown G. and Kuncheva L., "'Good' and 'Bad' Diversity in Majority Vote Ensembles," in *Proceedings of International Workshop on Multiple Classifier Systems*, Cairo, pp. 124-133, 2010.
- [4] Brzezinski D. and Stefanowski J., "Reacting to Different Types of Concept Drift: The Accuracy Updated Ensemble Algorithm," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 1, pp. 81-94, 2014.
- [5] Brzezinski D., Stefanowski J., Susmaga R. and Szczech I., "On the Dynamics of Classification Measures for Imbalanced and Streaming Data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 8, pp. 2868-2878, 2020.
- [6] Cano A. and Krawczyk B., "Kappa Updated Ensemble for Drifting Data Stream Mining," *Machine Learning*, vol. 109, no. 1, pp.175-218, 2020.
- [7] Ditzler G., Roveri M., Alippi C., and Polikar R., "Learning in Nonstationary Environments: A Survey," *IEEE Computational Intelligence Magazine*, vol. 10, no. 4, pp. 12-25, 2015.
- [8] Gama J., *Knowledge Discovery from Data Streams*, New York: CRC Press, 2010.
- [9] Gama J., Žliobaitė I., Bifet A., Pechenizkiy M., and Bouchachia A., "A Survey on Concept Drift Adaptation," *ACM Computing Surveys*, vol. 46, no. 4, pp. 231-238, 2014.
- [10] Gomes H., Barddal J., and Enembreck F., and

- Bifet A., "A Survey on Ensemble Learning for Data Stream Classification," *ACM Computing Surveys*, vol. 50, no. 2, pp.1-36, 2017.
- [11] Gomes H., Bifet A., and Read J., Barddal J., Enembreck F., Pfharinger B., Holmes G., and Abdessalem T., "Adaptive Random Forests for Evolving Data Stream Classification," *Machine Learning*, vol. 106, no. 9-10, pp. 1469-1495, 2017.
- [12] Gomes H., Read J., and Bifet A., Barddal J., and Gama J., "Machine Learning for Streaming Data: State of the Art, Challenges, and Opportunities," *ACM SIGKDD Explorations Newsletter*, vol. 21, no. 2, pp. 6-22, 2019.
- [13] Khamassi I., Sayed-Mouchaweh M., Hammami M., and Ghédira J., "Discussion and Review on Evolving Data Streams and Concept Drift Adapting," *Evolving Systems*, vol. 9, no. 1, pp. 1-23, 2018.
- [14] Liu A., Lu J., and Zhang G., "Diverse Instance-Weighting Ensemble Based on Region Drift Disagreement for Concept Drift Adaptation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 293-307, 2021.
- [15] Lu J., Liu A., Dong F., Gu F., Gama J., and Zhang G., "Learning under Concept Drift: A Review," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 12, pp. 2346-2363, 2019.
- [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] Morales G., Bifet A., Khan L., Gama J., and Fan W., "IoT Big Data Stream Mining," in *Proceedings of the 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, pp. 2119-2120, 2016.
- [18] Pietruczuk L., Rutkowski L., Jaworski M., and Duda P., "How to Adjust an Ensemble Size in Stream Data Mining?" *Information Sciences*, vol. 381, pp. 46-54, 2017.
- [19] Rijn J., Holmes G., and Pfahringer B., and Vanschoren J., "Having a Blast: Meta-Learning and Heterogeneous Ensembles for Data Streams," in *Proceedings of 15<sup>th</sup> International Conference on Data Mining*, Atlantic City, pp. 1003-1008, 2015.
- [20] Santos S., Jr P., and Silva G., and De Barros R., "Speeding Up Recovery from Concept Drifts," in *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases*, Nancy, pp. 179-194, 2014.
- [21] Street W. and Kim Y., "A Streaming Ensemble Algorithm (SEA) for Large-scale Classification," in *Proceedings of the 7<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, pp. 377-382, 2001.
- [22] Sun B., Wang J., and Chen H., and Wang Y., "Diversity Measurement in Ensemble Learning," *Control and Decision*, vol. 29, no. 3, pp. 385-395, 2014.
- [23] Sun Y., Tang K., and Zhu Z., and Yao X., "Concept Drift Adaptation by Exploiting Historical Knowledge," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 99, no. 10, pp. 1-11, 2017.
- [24] Sun Y., Wang Z., and Yuan J., and Zhang W., "Tracking Recurring Concepts from Evolving Data Streams using Ensemble Method," *The International Arab Journal of Information Technology*, vol. 16, no. 6, pp. 1044-1052, 2019.
- [25] Wang H., Fan W., and Yu P., and Han J., "Mining Concept-drifting Data Streams Using Ensembles Classifiers," in *Proceedings of the 9<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, pp. 226-235, 2003.
- [26] Webb G., Hyde R., and Cao H., Nguyen H., and Petitjean F., "Characterizing Concept Drift," *Data Mining and Knowledge Discovery*, vol. 30, no. 4, pp. 964-994, 2016.



**Yange Sun** is an associate professor at Xinyang Normal University, Xinyang, China. She received the Ph.D. degree in computer science and technology in 2019 from Beijing Jiaotong University Beijing, China and the M.S. degree from the Central China Normal University, in 2007, both in computer science. Her research interests include data mining and machine learning.



**Han Shao** is a lecture at Xinyang Normal University, Xinyang, China. His research interests include big data mining and machine learning.



**Bencai Zhang** born in 1994, Master. His current research interests include data mining and machine learning.