The Value of Diversity for Dealing with Concept Drift in Class Imbalanced Data Streams - Supplementary Document

Chun Wai Chiu, Leandro L. Minku, *Senior Member, IEEE,*

I. General

This is a supplementary document to the paper of "On the Benefit of Diversity for Dealing with Concept Drift in Class Imbalanced Data Streams" and it is organised as follows.

- Table I presents a list of symbols used in the paper.
- Section II presents the details of artificial and real-world data streams.
- Section III presents the pseudo-code of Concept Drift handling based on Clustering in the Model Space for Class Imbalanced Learning (CDCMS.CIL) written in a lower level of abstraction.
- Section IV presents the additional results on the predictive performance of approaches on artificial data streams and real-world data streams. It complements Section V-C to V-E of the paper.

II. Data Streams

This section presents the summary of artificial and real-world data streams which helps to comprehend Section V-A to V-B in the paper.

A. Artificial Data Streams:

Table II provides a summary of the characteristics of the artificial data streams and Table III presents the percentage difference between the concepts to denote the severity of concept drift in the artificial data streams.

We adopted four different data stream generators from the Massive Online Analysis (MOA) framework [3] to simulate data streams. This enables us to specify the properties of the drifts, which help us to understand the effectiveness of data stream learners in dealing with various types of concept drift with different severity, speed and recurrent behaviour in class-imbalanced data streams. The generators we adopted were the Sine [4], Agrawal [5], SEA [6], and STAGGER [7].

The Sine generator [4] has four numerical input attributes (two are irrelevant) and uses different mathematical sine functions as the decision boundaries for different concepts. The Agrawal generator [5] has nine input attributes. The classification task is to predict whether a loan should be granted to an individual. The SEA generator [6] has three numerical input attributes (one is irrelevant). The decision boundary is a hyperplane with different thresholds (8, 9, 7 and 9.5) to represent different concepts. The STAGGER generator [7] has three categorical attributes (size, colour and shape). Depending on the concept, only one or more input attributes are actually used in the classification. Each generator has a set of functions {f1,...,fn} to produce different concepts. An additional SEA concept f5 with threshold of 4 is included in our study to enable a relatively higher drift severity.

Concept drifts were simulated using a sigmoid function to connect two concepts and to decide the probability of the data coming from the pre- and post-drift concepts. Each concept has a length of 20k time steps / examples. Table II summarises the details of the artificial data streams. Two versions of each stream were simulated – one containing abrupt and the other containing gradual drifts, which complete in one and 2k time steps, respectively. This leads to 20 data streams. Each drift contains a different level of severity as listed in our supplementary document.

We also enforced four different class-imbalance ratio on each of the 20 data streams: 7:3, 8:2 and 9:1. Thus, a total of $10 \times 2 \times 4 = 80$ artificial data streams were used in the experiments. All data streams are normalised.

B. Real-world Data Streams

Nine real-world data streams were adopted to evaluate the predictive performance of the data stream learners in practical applications. They were obtained from [8]. The Airline stream [9] consists of around 20 years commercial flight records of arrival and departure. The NOAA stream [10] contains weather measurements collected over five decades. The Forest Cover type (Covtype) streams [11] contains the information about the forest cover type of 30×30-meter cells. The Luxembourg stream was constructed using the European Social Survey 2002-2007, concerning internet usage [12]. The Ozone stream consists of air measurements from 1998 to 2004 at the Houston, Galveston, and Brazoria areas [13]. The PAKDD2009 stream consists of credit scoring data from the private label credit card operation of a major Brazilian retail chain [14]. The Amazon stream comprises labelled positive and negative reviews for books, DVDs, electronics, and kitchen appliances, obtained from 1998 to 2004 [15]. The Twitter stream contains annotated tweets collected from July to December 2015 [16]. The INSECTS streams [8] were constructed using a smart trap to collect the flying data of three different species of insects in a climate controlled environment. Note that, [8]

Chun Wai Chiu is with Baxall Construction Ltd, Paddock Wood, Kent, TN12 6BU, UK. E-mail: mchiu@baxallconstruction.co.uk

Leandro L. Minku is with the School of Computer Science, University of Birmingham, Birmingham, B15 2TT, UK. E-mail: L.L.Minku@bham.ac.uk

TABLE I
SYMBOLS IN THE PAPER AND THEIR DESCRIPTIONS

Symbol	Description
DS	Data Stream, a potentially infinite sequence of data.
(x_t, y_t)	Data (example) that arrives at time t. x_t is a d-dimensional feature vector. y_t is the target label.
D_t	$= P_t(x_t, y_t)$ Concept (the underlying joint probability distribution of data) at time t.
$P_t(y^i)$	The prior probability of class y^i at time t.
ϕ	Class size fading factor
λ	λ parameter of Poisson distribution
NH	The heterogeneously diverse auxiliary ensemble of CDCMS.CIL
NL	The new main ensemble of CDCMS.CIL after drift detection
c	A base learner of CDCMS.CIL
c_i	The i-th base learner of CDCMS.CIL
c_n	The new base learner of CDCMS.CIL Created upon drift detection
c_{worst}	The least performing base learner in the main ensemble
c_n^{memory}	The n-th base learner in the model repository (memory)
c_{sim}	The base learner in the model repository (memory) that is the most similar to c_{worst}
θ	Similarity threshold based on Yule's Q-Statistics [1]
$SC[]$	An array of stream clustering methods
mc	Micro-cluster, summaries n d-dimensional points.
$\overline{CF2^x}$	$= (\sum_{i=1}^{n} (x_{t_i,1})^2, \ldots, \sum_{i=1}^{n} (x_{t_i,d})^2)$, The squared sum of features in mc
$\overline{CF1^x}$	$= (\sum_{i=1}^{n} x_{t_i,1}, \ldots, \sum_{i=1}^{n} x_{t_i,d})$, The linear sum of features in mc
$CF2^t$	$= \sum_{i=1}^{n} (t_i)^2$, The squared sum of time stamps in mc
$CF1^t$	$= \sum_{i=1}^{n} t_i$, The linear sum of time stamps in mc
n	The number of examples (data) that the mc summaries
$MC_{c_i}^{class}$	A set of micro-clusters associated to c_i for $class$
K_i	The number of mc in $MC_{c_i}^{class}$
S	The overall sparse representation of past data Consist of all mc in CDCMS.CIL
$\overline{mc}$	The centre of a micro-cluster
$\overline{MC}$	A set of micro-clusters' centres
P	Project set Consist of all micro-clusters' centres in CDCMS.CIL
P'	Oversampled projection set by round robin method
$RR()$	Round Robin Method
C	Clustering examples A set of prediction correctness on P' by a given base learner
$\mathbf{z}$	$\in \{0, 1\}$, Prediction correctness
L	Number of base learners
$Q()$	Yule's Q-Statistics [1]
$N^{a,b}$	Count of prediction where c_i is a and c_j is b, for any $i, j \le L$
Q_{all}	$= \{Q(c_{worst}, c_1^{memory}), \ldots, Q(c_{worst}, c_n^{memory})\}$ A set of pairwise Q-Statistics between c_{worst} and each base learner in the model repository (memory)

TABLE II
SUMMARY OF ARTIFICIAL DATA STREAMS (TABLE FROM [2])

Stream	#Instances	#Nom. Attr.	#Num Attr.	Concept Sequence	IR
Sine1	60k	0	4	f3→f4→f3	{0.7:0.3 / 0.8:0.2 / 0.9:0.1}
Sine2	100k	0	4	f1→f2→f3→f4→f1	
Agr1	100k	3	6	f1→f3→f4→f7→f10	
Agr2	120k	3	6	f7→f4→f6→f5→f2→f9	
Agr3	120k	3	6	f4→f2→f1→f3→f4	
Agr4	100k	3	6	f1→f3→f6→f5→f4	
SEA1	100k	0	3	f5→f3→f1→f2→f4	
SEA2	100k	0	3	f5→f1→f4→f3→f2	
STA1	80k	3	0	f1→f2→f3→f2	
STA2	80k	3	0	f2→f3→f1→f2	

- Agr: Agrawal; STA: STAGGER; Nom. Attr.: Nominal attributes; Num. Attr.: Numeric attributes; IR: Imbalance Ratio
- Lime (light grey) rows refers to as data streams presenting recurring concepts.
- fn represents the n-th function of the stream type, i.e., f1 in Sine1 refers to the first function of the Sine generator.

TABLE III
PERCENTAGE DIFFERENCE OF CONCEPTS

	Sine			SEA				STAGGER	
	f1	f2	f3	f1	f2	f3	f4	f1	f2
f2	100.0%	-	-	8.5%	-	-	-	59.3%	-
f3	26.8%	73.2%	-	7.4%	16.0%	-	-	77.8%	48.1%
f4	73.2%	26.8%	100.0%	13.1%	4.6%	20.6%	-	-	-
f5	-	-	-	23.9%	32.5%	16.5%	37.1%	-	-

Agrawal

	f1	f2	f3	f4	f5	f6	f7	f8	f9
f2	53.9%	-	-	-	-	-	-	-	-
f3	53.1%	50.8%	-	-	-	-	-	-	-
f4	53.9%	20.5%	50.8%	-	-	-	-	-	-
f5	53.4%	47.6%	50.7%	47.7%	-	-	-	-	-
f6	69.9%	28.9%	51.2%	35.5%	48.1%	-	-	-	-
f7	50.5%	53.3%	50.1%	53.5%	60.1%	57.2%	-	-	-
f8	33.5%	60.4%	46.5%	59.6%	59.6%	59.8%	49.8%	-	-
f9	50.4%	53.3%	50.2%	53.5%	59.9%	57.3%	6.0%	49.5%	-
f10	32.9%	61.3%	46.5%	61.3%	60.0%	59.9%	51.1%	1.8%	51.1%

All percentage differences were calculated based on one million random generated examples.

proposed seven INSECTS streams but we only adopted six of them which contain concept drifts and left the INSECT-out-of-control stream unused as it does not contain any concept drift [8].

Covtype and INSECTS were originally multi-class problems. To adapt them for our study, we converted them into several versions of binary classification problems. The binary attributes of Covtype were first converted to nominal format. Covtype was converted into seven different binary classification problems, each taking one of the seven classes as class 1 while combining the other classes as class 0. This allow us to evaluate how well does a given data stream learner in learning each of the seven classes.

The INSECTS streams originally comprised six classes: three species of mosquitoes with two genders. We transformed them into binary classification problems by considering classes belonging to the species of ae-albopictus as the minority class and combining the remaining classes as the majority class.

Table IV summaries the details of the real-world data streams.

III. PROPOSED APPROACH

Algorithm 1 presents the pseudo-code, over-viewing CDCMS.CIL. The highlighted code fragments contrast the differences between CDCMS and CDCMS.CIL. Section III-A presents the details of the augmentation done to base learners facilitating the proposed diversity-based memory strategies for class imbalanced learning. Section III-B presents the nominal feature handling for stream clustering.

A. Augmented Base Learner

Algorithm 2 presents the training procedure of an augmented base learner. It first trains the actual base learner

Algorithm 1 Concept Drift Handling Based on Clustering in the Model Space for Class Imbalanced Learning - CDCMS.CIL

Hyper-parameters: Base learner (c_{type}), Ensemble Size (e), Memory Size ($e \times r$), Similarity Threshold (θ), Stream Clustering Method (sc), Class Size Fading Factor(α), Weighting Metric (w), Resampling Method ($resampler$), Data Stream (S)

Variables: New Ensemble (E_{nl}), Old Ensemble (E_{ol}), Highly Diverse Ensemble (E_{nh}) New Model (c_n), Memory (R), Clustering Result (C), Cluster ($cluster$), Stream Clustering Methods array ($SC[]$)

$SC[] \leftarrow createStreamClusteringMethods(sc, 2)$ ▷ "2" refers to the number of SC to create.
$E_{nl} \leftarrow E_{nl}.add(createLearnerWithInfo(c_{type}, resampler, \alpha, SC[].copy()))$
$E_{ol} \leftarrow null$
$E_{nh} \leftarrow null$
$c_n \leftarrow createLearnerWithInfo(c_{type}, resampler, \alpha, SC[].copy())$
for $s_t \in S$ **do**
  $drift\_level \leftarrow DriftDetection(E_{nl}, s_t)$
  if $drift\_level == NORMAL$ **then**
    if t mod $b == 0$ **then**
      if t is b time steps after a drift **then**
        $C \leftarrow (R \cup c_n).clusteringModelsBasedOnMicroClusters()$
        $cluster \leftarrow C.getCluster(c_n)$
        if cluster.size() == 1 **then**
          $E_{nl}.add(c_n)$
        else
          $E_{nl}.add(cluster.getMostTrained(e-1))$
        end if
      else if $|E_{nl}| \geq e$ **then**
        $c_{worst} \leftarrow E_{nl}.getWorstModel()$
        $c_{worst}.resetWeight()$
        $R.saveByDiversityBasedOnMicroClusters(c_{worst}, \theta)$
      end if
      $c_n \leftarrow createLearnerWithInfo(c_{type}, resampler, \alpha, SC[].copy())$
      $E_{nl}.add(c_n)$
    end if
    $c_n.updateWeight(s_t)$
    $c_n.trainOnInstance(s_t)$
    $drift\_level_{previous} \leftarrow NORMAL$
  else if $drift\_level == DRIFT$ **then**
    $E_{ol} \leftarrow E_{nl}$
    $E_{nl}.resetWeight()$
    $R.saveByDiversityBasedOnMicroClusters(E_{nl})$
    $E_{nl}.clear()$
    $E_{nh} \leftarrow createEnsemble(e)$
    if $drift\_level_{previous} == NORMAL$ **then**
      $c_n \leftarrow createLearnerWithInfo(c_{type}, resampler, \alpha, SC[].copy())$
      $C \leftarrow R.clusteringModelsBasedOnMicroClusters()$
      $E_{nh} \leftarrow C.getRepresentativeModels()$
    end if
    $E_{nl}.add(c_n)$
    $c_n \leftarrow createLearnerWithInfo(c_{type}, resampler, \alpha, SC[].copy())$
    $E_{nh}.resetWeight(); E_{nl}.resetWeight(); E_{ol}.resetWeight()$
    $drift\_level_{previous} \leftarrow DRIFT$
  end if
  if $E_{ol} \neq null$ **then**
    $E_{ol}.updateWeight(s_t)$
  end if
  if $E_{nh} \neq null$ **then**
    $E_{nh}.updateWeight(s_t)$
  end if
  $E_{nl}.updateWeight(s_t)$
  $E_{nl}.trainOnInstance(s_t)$
end for

with s_t (line 2, Algorithm 2). After that, it makes a copy of s_t and prepares it to be ready for training the stream clustering method (line 3, Algorithm 2). We denotes the copy of s_t to as $s_t^{Bin,noClass}$. As most stream clustering methods cannot handle nominal attributes directly and they should not take the class attribute into the account when creating and managing the micro-clusters, all the nominal attributes in $s_t^{Bin,noClass}$ have to be converted to binary attributes using one-hot encoding, and the class attribute has to be removed (line 3-4, Algorithm 2). Lastly, the stream clustering method, which corresponds to the class value of s_t, trains on $s_t^{Bin,noClass}$ (line 5, Algorithm 2).

To reduce the learning bias towards the majority class, CDCMS.CIL employs a resampling method to train the main ensemble E_{nl}. This resampling method can also be any method from the literature but, for simplicity, only simple

TABLE IV
SUMMARY OF REAL-WORLD DATA STREAMS

Stream	#Inst.	#Nom. Attr.	#Num. Attr.	Imbalance Ratio
Airlines	485,445	4	3	0.544:0.486
NOAA	16,344	0	8	0.685:0.315
Covtype$_{(c_1=\{1\text{-}6\})}$	522,911	2	10	0.619:0.381
Covtype$_{(c_1=1)}$	522,911	2	10	0.524:0.476
Covtype$_{(c_1=2)}$	522,911	2	10	0.936:0.064
Covtype$_{(c_1=3)}$	522,911	2	10	0.999:0.001
Covtype$_{(c_1=4)}$	522,911	2	10	0.987:0.014
Covtype$_{(c_1=5)}$	522,911	2	10	0.971:0.029
Covtype$_{(c_1=6)}$	522,911	2	10	0.965:0.035
INSECTS$^{\text{inc.}}$	406,840	0	33	0.905:0.095
INSECTS$_{\text{abr.}}$	319,748	0	33	0.907:0.093
INSECTS$^{\text{inc.}}_{\text{grad.}}$	128,981	0	33	0.899:0.101
INSECTS$^{\text{inc.}}_{\text{abr. re.}}$	406,840	0	33	0.905:0.095
INSECTS$^{\text{inc.}}_{\text{re.}}$	406,840	0	33	0.905:0.095
Luxembourg	1711	15	16	0.512:0.488
Ozone	2,281	0	72	0.942:0.058
PAKDD-2009	44998	13	14	0.803:0.197
Amazon	7,200	0	30	0.875:0.125
Twitter	8,181	0	30	0.846:0.154

- Inst: Instances; Nom. Attr.: Nominal attributes; Num. Attr.: Numeric attributes
- Total number of attributes = #Nom. Attr. + #Num. Attr + Class attribute.
- Covtype$_{(c_1=x)}$: "c$_1$=x" denotes class 1 as x in the original dataset, with the remaining classes combined as class 0. Similarly, "c$_1$=$x_0 - x_n$" represents class 1 as x_0-x_n combined in the original dataset, with the remaining classes combined as class 0.
- INSECTS streams: "ae-albopictus" is the class 1; "inc.": incremental; "abr.": abrupt, "grad": gradual, and "re.": recurring.

Algorithm 2 Train LearnerWithInfo on Instance s_t

```
function TRAINONINSTANCE(s_t)
    base_learner.trainOnInstance(s_t)
    s_t^Bin ← nominalToBinary(s_t.copy())
    s_t^{Bin,noClass} ← s_t^Bin.deleteClassAttribute()
    SC[s_t.classValue()].trainOnInstance(s_t^{Bin,noClass})
end function
```

oversampling, simple undersampling and no resampling are considered in this study. The procedure of training the ensembles is the same as CDCMS, except the most recent training instance is weighted according to the current estimated class size. Eq. 1 presents the class size estimation at the current time step t.

$$classSize(k)_t = \theta \times classSize(k)_{(t-1)} + [(s_t, c_k)] \quad (1)$$

where $classSize(k)_t$ refers to the class size of the k-th class at time step t. $[(s_t, c_k)] = 1$ if the true class label of s_t is c_k, otherwise 0. θ ($0 < \theta < 1$) is a predefined time decaying factor to make the past examples less important to the current class size. When simple oversampling (undersampling) is used, the weight of minority (majority) class examples scales by $\frac{classSize(Maj)_t}{classSize(Min)_t}$ ($\frac{classSize(Min)_t}{classSize(Maj)_t}$), whilst the weight of majority (minority) class examples remains unchanged.

B. Stream Clustering - Nominal Feature Handling

Typically, stream clustering methods require one-hot encoding of nominal features. Therefore, the synthetic examples in P will normally use one-hot encoding. However, popular online base learning algorithms such as Hoeffding Trees [17] can work well without one-hot encoding. In this case, we convert P back to the original feature space to enable predictions by the base learners. To convert a set of dummy features $G = \{d_1, d_2, \ldots, d_k\}$, which corresponds to a set of possible categorical values $V = \{v_1, v_2, \ldots, v_k\}$ of a nominal feature F_{nom}, we perform the following:

$$F_{nom} = \arg\max(G)$$

This operation maps G back to V and assigns F_{nom} with v_i, where $1 \leq i \leq k$, corresponding to the highest value in G.

IV. ADDITIONAL RESULTS

This section presents additional results in the forms of tables and figures to complement Section V-C to V-E. Note that plots are based on the first run of the approaches on each stream.

- Table V presents the Friedman ranks of CDCMS.CIL with different weighting metrics and resampling strategies in terms of time-decay G-Mean, majority class recall, and minority class recall.
- Table VI presents the performance values of CDCMS.CIL with different weighting metrics and resampling strategies in terms of time-decay G-Mean, majority class recall, and minority class recall.
- Figures 1-10 compare the time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on artificial data streams.
- Figures 21-39 compare the predictive performance (time-decay G-Mean, prequential accuracy, class 0 and class 1 recalls) of CDCMS.CIL with different weighting metrics and resampling strategies on real-world data streams.
- Figures 11-20 compare the time-decay G-Mean of homogeneous and heterogeneous diversity approaches on artificial data streams.
- Figures 21-39 compare the predictive performance (time-decay G-Mean, prequential accuracy, class 0 and class 1 recalls) of homogeneous and heterogeneous diversity approaches on real-world data streams.

TABLE V
FRIEDMAN RANKS OF CDCMS.CIL WITH DIFFERENT WEIGHTING METRICS AND RESAMPLING STRATEGIES IN TIME-DECAY G-MEAN, CLASS 0, AND CLASS 1 RECALLS

Groups	Preq. Acc.	Preq. Acc. OS	Preq. Acc. US	GMean	GMean OS	GMean US
Time-Decay G-Mean						
Grouped by imbalance ratio (Artificial data streams)						
0.5:0.5	3.47	2.56	5.01	3.22	2.47	4.28
0.7:0.3	5.38	2.38	3.18	5.17	2.26	2.64
0.8:0.2	5.39	2.31	3.12	5.26	2.28	2.64
0.9:0.1	5.39	2.37	3.00	5.07	2.70	2.48
Grouped by concept drift speed						
Abr.	4.82	2.12	3.80	4.72	2.27	3.27
Grad.	4.99	2.69	3.35	4.63	2.59	2.75
Real	3.89	3.16	4.37	2.89	2.63	4.05
Grouped by streams						
Sine	5.04	1.96	4.15	4.86	1.67	3.31
Agr.	4.61	2.18	3.33	4.66	3.05	3.16
SEA	5.73	3.19	4.00	4.29	1.72	2.06
STA.	4.54	2.48	3.06	4.92	2.65	3.34
Airlines	6.0	4.0	1.0	5.0	2.0	3.0
NOAA	6.0	4.0	3.0	5.0	1.0	2.0
Covtype	2.86	2.29	4.43	3.71	2.86	4.86
INSECTS	3.60	4.00	5.20	1.20	2.60	4.40
Luxembourg	6.0	4.0	2.0	1.0	5.0	3.0
Ozone	1.0	6.0	5.0	4.0	2.0	3.0
PAKDD-2009	6.0	4.0	5.0	1.0	3.0	2.0
Amazon	6.0	1.0	4.0	3.0	2.0	5.0
Twitter	5.0	1.0	6.0	4.0	2.0	3.0
All	4.71	2.55	3.73	4.34	2.47	3.21
Time-Decay Class 0 (Majority Class) Recall						
Grouped by imbalance ratio (Artificial data streams)						
0.5:0.5	3.96	3.12	4.27	3.55	2.81	3.29
0.7:0.3	1.38	3.09	4.92	2.09	3.84	5.69
0.8:0.2	1.30	3.08	4.68	2.16	3.93	5.85
0.9:0.1	1.27	2.93	4.80	2.22	3.91	5.86
Grouped by concept drift speed						
Abr.	2.18	2.88	4.69	2.66	3.42	5.18
Grad.	1.78	3.24	4.65	2.35	3.82	5.16
Real	2.26	2.74	4.00	2.79	3.89	5.32
Grouped by streams						
Sine	1.67	2.98	5.01	2.34	3.38	5.61
Agr.	1.59	2.90	5.06	2.38	3.80	5.28
SEA	2.15	3.97	3.05	2.37	4.65	4.80
STA.	2.89	2.53	5.15	3.05	2.49	4.88
Airlines	1.0	3.0	5.0	2.0	4.0	6.0
NOAA	1.0	3.0	4.0	2.0	6.0	5.0
Covtype	2.29	2.86	4.29	3.29	2.86	5.43
INSECTS	2.20	1.80	5.20	2.60	3.40	5.80
Luxembourg	6.0	5.0	2.0	1.0	4.0	3.0
Ozone	5.0	2.0	1.0	3.0	6.0	4.0
PAKDD-2009	1.0	3.0	2.0	4.0	6.0	5.0
Amazon	1.0	3.0	4.0	2.0	6.0	5.0
Twitter	1.0	4.0	2.0	3.0	5.0	6.0
All	2.03	3.00	4.54	2.56	3.68	5.20
Time-Decay Class 1 (Minority Class) Recall						
Grouped by imbalance ratio (Artificial data streams)						
0.5:0.5	2.79	2.29	4.83	3.40	2.83	4.86
0.7:0.3	5.61	3.27	2.45	5.26	2.82	1.59
0.8:0.2	5.60	2.89	2.86	5.31	2.57	1.77
0.9:0.1	5.57	2.69	2.79	5.16	2.92	1.87
Grouped by concept drift speed						
Abr.	4.85	2.50	3.43	4.85	2.62	2.75
Grad.	4.94	3.07	3.04	4.71	2.96	2.30
Real	4.37	3.79	3.74	3.42	2.74	2.95
Grouped by streams						
Sine	5.15	2.94	3.51	4.72	2.49	2.19
Agr.	4.81	2.77	3.00	4.73	3.11	2.57
SEA	5.02	2.85	4.14	4.58	2.22	2.19
STA.	4.67	2.58	2.51	5.14	3.00	3.10
Airlines	6.0	4.0	2.0	5.0	3.0	1.0
NOAA	6.0	4.0	3.0	5.0	1.0	2.0
Covtype	3.0	2.71	3.86	3.71	3.0	4.71
INSECTS	5.20	5.80	3.40	2.80	2.80	1.0
Luxembourg	5.0	4.0	2.0	1.0	6.0	3.0
Ozone	2.0	6.0	5.0	4.0	1.0	3.0
PAKDD-2009	6.0	4.0	5.0	3.0	2.0	1.0
Amazon	6.0	1.0	4.0	3.0	2.0	5.0
Twitter	5.0	1.0	6.0	4.0	2.0	3.0
All	4.79	2.98	3.33	4.52	2.78	2.61

- The p-values of Friedman tests are all $\leq$2.2E-16.
- Highlighted ranks denote significant superior performance.
- "Abr": Aburpt; "Grad.": Gradual; "Agr.": Agrawal; "STA.": STAGGER; "Real": Real-world

TABLE VI
MEDIAN OF PREDICTIVE PERFORMANCE OF CDCMS.CIL WITH DIFFERENT WEIGHTING METRICS AND RESAMPLING STRATEGIES IN TIME-DECAY G-MEAN, CLASS 0, AND CLASS 1 RECALLS

Groups	Preq. Acc.	Preq. Acc. OS	Preq. Acc. US	GMean	GMean OS	GMean US
Time-Decay G-Mean						
Grouped by imbalance ratio (Artificial data streams)						
0.5:0.5	86.33%	86.32%	85.47%	86.02%	85.99%	85.24%
0.7:0.3	80.98%	85.02%	84.24%	80.97%	84.56%	84.49%
0.8:0.2	74.82%	81.18%	80.58%	74.82%	80.55%	80.77%
0.9:0.1	64.94%	72.87%	70.55%	64.68%	72.78%	77.68%
Grouped by concept drift speed						
Abr.	81.01%	83.54%	82.36%	81.11%	83.34%	82.93%
Grad.	80.34%	82.35%	81.74%	80.23%	82.53%	82.26%
Real	66.86%	68.02%	67.54%	67.48%	69.68%	68.35%
Grouped by streams						
Sine	84.68%	89.98%	89.42%	84.81%	90.05%	89.5%
Agr.	76.24%	79.9%	79.22%	76.36%	78.9%	79.33%
SEA	62.46%	78.31%	77.61%	66.57%	79.81%	79.82%
STA.	98.8%	99.06%	99.04%	98.8%	99.03%	99.01%
Airlines	60.75%	62.02%	63.92%	62.0%	63.67%	63.59%
NOAA	66.86%	68.02%	68.22%	67.48%	70.8%	68.35%
Covtype	64.26%	64.86%	62.12%	62.55%	64.95%	60.58%
INSECTS	70.28%	69.77%	68.84%	72.79%	70.54%	69.52%
Luxembourg	84.3%	86.41%	91.56%	93.74%	84.71%	91.16%
Ozone	66.45%	55.88%	57.18%	60.4%	65.24%	64.01%
PAKDD-2009	41.78%	45.48%	44.7%	56.63%	53.99%	56.27%
Amazon	45.39%	54.29%	48.53%	50.52%	51.8%	48.19%
Twitter	55.23%	61.36%	52.61%	56.75%	58.79%	58.04%
All	78.42%	81.02%	80.86%	78.46%	81.02%	80.95%
Time-Decay Class 0 (Majority Class) Recall						
Grouped by imbalance ratio (Artificial data streams)						
0.5:0.5	86.95%	87.1%	86.7%	86.82%	86.98%	87.04%
0.7:0.3	95.62%	92.07%	88.54%	94.95%	91.2%	88.2%
0.8:0.2	97.69%	94.1%	90.27%	96.74%	92.42%	88.29%
0.9:0.1	99.44%	96.64%	92.56%	98.14%	94.1%	88.01%
Grouped by concept drift speed						
Abr.	97.78%	94.09%	89.99%	96.88%	92.13%	88.49%
Grad.	97.51%	93.46%	89.58%	96.55%	91.55%	87.28%
Real	88.1%	85.85%	87.8%	86.33%	83.46%	83.99%
Grouped by streams						
Sine	98.46%	95.56%	93.17%	97.04%	95.46%	92.39%
Agr.	96.39%	93.16%	88.66%	95.53%	90.86%	86.08%
SEA	90.33%	81.43%	81.82%	87.68%	77.7%	75.67%
STA.	99.79%	99.72%	99.22%	99.76%	99.66%	99.25%
Airlines	76.71%	70.51%	66.13%	74.96%	69.84%	64.92%
NOAA	79.67%	78.48%	76.85%	78.74%	74.14%	75.99%
Covtype	98.92%	98.82%	97.95%	99.15%	99.17%	98.67%
INSECTS	73.94%	78.64%	67.35%	75.14%	70.6%	65.73%
Luxembourg	88.1%	89.57%	95.36%	96.04%	89.93%	95.15%
Ozone	80.88%	91.27%	92.95%	86.89%	78.19%	86.1%
PAKDD-2009	84.5%	78.59%	80.09%	58.97%	48.12%	48.72%
Amazon	88.74%	85.32%	84.93%	85.97%	83.46%	84.86%
Twitter	89.54%	85.85%	89.11%	86.33%	84.82%	83.99%
All	97.21%	93.19%	89.57%	96.22%	91.05%	87.15%
Time-Decay Class 1 (Minority Class) Recall						
Grouped by imbalance ratio (Artificial data streams)						
0.5:0.5	90.01%	90.31%	88.96%	89.64%	89.93%	88.88%
0.7:0.3	68.76%	83.45%	81.97%	69.75%	84.9%	86.51%
0.8:0.2	57.91%	74.27%	72.69%	58.65%	80.3%	83.27%
0.9:0.1	45.01%	56.67%	55.15%	44.42%	63.5%	78.07%
Grouped by concept drift speed						
Abr.	72.39%	81.71%	80.59%	73.66%	83.86%	85.48%
Grad.	69.24%	80.35%	79.79%	70.29%	81.87%	84.37%
Real	62.33%	59.84%	62.06%	58.14%	66.32%	62.59%
Grouped by streams						
Sine	73.7%	87.65%	87.55%	74.61%	88.03%	88.79%
Agr.	61.8%	70.08%	72.07%	62.67%	69.32%	73.75%
SEA	49.51%	75.54%	73.77%	53.66%	82.89%	84.93%
STA.	98.48%	98.84%	98.94%	98.46%	98.78%	98.85%
Airlines	51.5%	56.07%	62.06%	53.89%	58.8%	62.41%
NOAA	56.48%	59.37%	60.92%	58.14%	68.09%	61.78%
Covtype	58.85%	59.84%	56.82%	55.77%	60.05%	52.92%
INSECTS	66.74%	65.51%	70.67%	70.89%	71.02%	73.8%
Luxembourg	80.92%	83.61%	88.84%	92.28%	80.46%	88.42%
Ozone	62.33%	38.22%	39.25%	47.56%	62.92%	54.5%
PAKDD-2009	21.72%	27.12%	26.29%	56.65%	63.15%	67.36%
Amazon	23.31%	35.17%	28.05%	30.35%	33.05%	27.93%
Twitter	34.43%	44.37%	32.78%	38.18%	41.45%	40.98%
All	66.74%	78.47%	78.26%	69.53%	80.25%	80.7%

- The p-values of Friedman tests are all $\leq$2.2E-16.
- Highlighted ranks denote significant superior performance.
- "Abr": Aburpt; "Grad.": Gradual; "Agr.": Agrawal; "STA.": STAGGER; "Real": Real-world

TABLE VII
FRIEDMAN RANKS OF APPROACHES IN TIME-DECAY G-MEAN, CLASS 0, AND CLASS 1 RECALLS

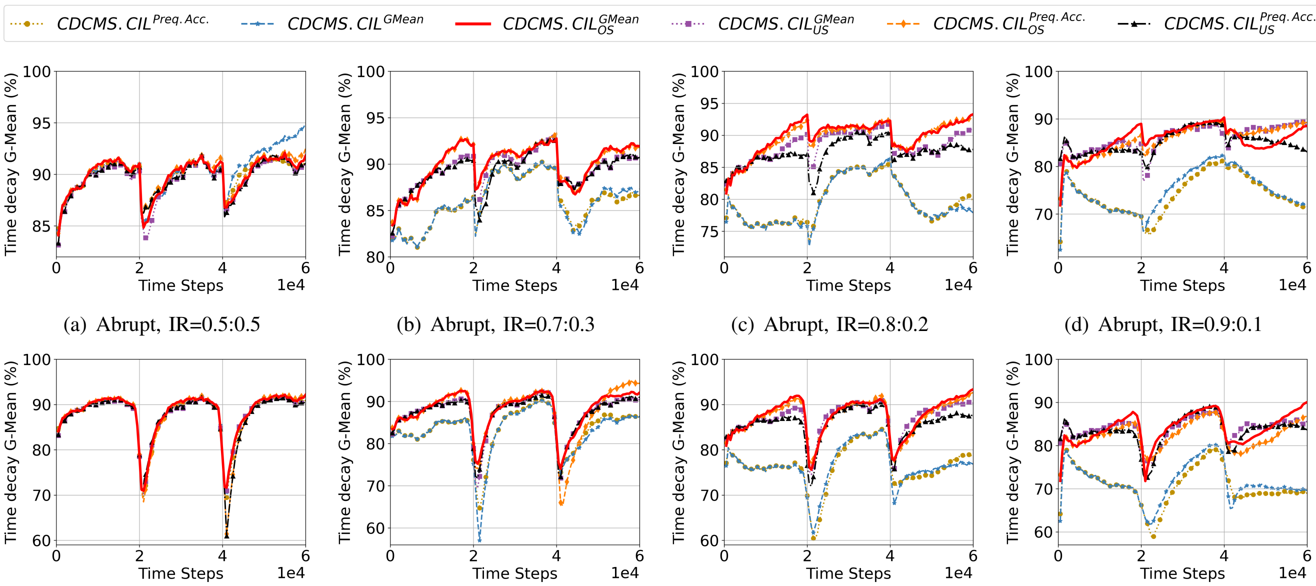
Groups	GH-VFDT$_d$	HD-VFDT$_d$	Oza-Bag$_d$	OOB$_d$	UOB$_d$	CSARF	VFC-SMOTE	SMOTE-OB	CDCMS.CIL$_{OS}^{GMean}$
Time-Decay G-Mean									
Grouped by imbalance ratio (Artificial data streams)									
0.5:0.5	5.90	5.98	3.23	3.99	5.04	4.42	8.16	6.18	2.12
0.7:0.3	6.49	6.65	6.33	3.15	3.84	4.53	7.45	4.11	2.44
0.8:0.2	6.88	7.36	6.84	2.74	2.93	4.77	7.37	2.93	3.18
0.9:0.1	7.25	7.64	6.23	3.19	2.40	4.47	7.26	2.48	4.09
Grouped by concept drift speed									
Abr.	6.49	6.78	5.34	3.23	3.71	4.83	7.44	4.51	2.67
Grad.	6.77	7.04	5.98	3.30	3.39	4.26	7.68	3.34	3.25
Real	5.74	5.79	6.10	4.39	5.51	1.78	8.54	2.24	4.92
Grouped by streams									
Sine	7.99	8.27	6.18	2.46	3.83	2.21	6.40	4.68	2.97
Agr.	5.73	5.67	4.78	4.15	3.83	6.38	8.94	2.52	2.99
SEA	7.57	7.62	7.34	2.64	3.53	4.46	6.72	2.03	3.09
STA.	6.13	7.30	5.21	2.91	2.74	3.30	6.79	7.88	2.75
Airlines	6.00	7.10	7.93	4.70	4.30	2.30	8.97	2.70	1.00
NOAA	6.07	6.07	7.93	3.97	5.73	1.00	9.00	2.57	2.67
Covtype	3.89	3.88	6.47	5.36	6.97	1.00	8.40	2.57	6.47
INSECTS	7.31	7.20	5.12	3.82	4.79	2.32	9.00	1.39	4.04
Luxembourg	7.47	7.47	2.80	2.37	6.03	1.00	4.37	4.50	9.00
Ozone	7.50	7.50	3.57	2.93	5.73	1.00	9.00	2.93	4.83
PAKDD-2009	6.50	6.50	8.00	4.80	4.20	1.00	9.00	2.37	2.63
Amazon	6.07	4.93	7.20	2.90	3.07	7.33	9.00	1.00	3.50
Twitter	5.57	7.40	7.60	5.10	2.77	1.67	9.00	1.60	4.30
All	6.46	6.69	5.74	3.48	3.93	4.02	7.75	3.60	3.33
Time-Decay Class 0 (Majority Class) Recall									
Grouped by imbalance ratio (Artificial data streams)									
0.5:0.5	5.82	5.92	3.84	4.77	5.10	3.71	7.73	5.98	2.13
0.7:0.3	3.76	3.49	1.56	6.32	8.42	3.97	4.47	8.15	4.87
0.8:0.2	3.74	3.53	1.78	6.25	8.58	4.31	3.35	8.29	5.18
0.9:0.1	3.30	3.24	2.01	6.05	8.55	4.83	3.15	8.44	5.44
Grouped by concept drift speed									
Abr.	4.29	4.24	2.34	5.82	7.53	3.89	4.95	7.99	3.95
Grad.	4.02	3.85	2.25	5.88	7.79	4.52	4.40	7.44	4.85
Real	4.39	4.67	3.73	4.89	7.03	5.79	2.53	7.00	4.97
Grouped by streams									
Sine	4.54	4.62	2.44	5.25	7.50	1.02	6.47	8.17	4.99
Agr.	4.40	3.75	1.72	6.33	8.00	5.09	3.68	7.06	4.97
SEA	2.73	2.79	2.63	6.62	8.13	5.35	5.42	7.44	3.89
STA.	4.71	5.31	2.97	4.71	6.67	4.47	4.12	8.84	3.19
Airlines	1.67	3.27	1.40	4.60	9.00	7.27	5.13	7.13	5.53
NOAA	7.97	7.97	2.00	4.87	4.33	3.63	1.00	8.03	5.20
Covtype	5.13	5.41	5.90	4.14	6.69	1.00	8.45	2.57	6.47
INSECTS	3.77	3.76	2.16	5.22	8.67	6.15	1.00	7.97	6.31
Luxembourg	7.47	7.47	2.68	2.45	4.53	1.03	5.62	4.78	8.97
Ozone	2.50	2.50	4.97	9.00	4.03	7.00	1.00	8.00	6.00
PAKDD-2009	3.50	3.50	2.00	5.10	5.90	7.00	1.00	9.00	8.00
Amazon	1.63	4.57	3.43	6.13	7.80	3.40	2.03	9.00	7.00
Twitter	4.00	2.77	2.23	5.60	7.83	9.00	1.00	7.13	5.43
All	4.20	4.16	2.57	5.66	7.54	4.51	4.26	7.58	4.51
Time-Decay Class 1 (Minority Class) Recall									
Grouped by imbalance ratio (Artificial data streams)									
0.5:0.5	5.36	5.44	3.33	3.99	4.86	5.01	7.61	6.30	3.11
0.7:0.3	6.72	7.02	6.77	3.59	1.66	5.22	7.43	3.10	3.50
0.8:0.2	6.92	7.43	6.93	3.20	1.64	5.20	7.38	2.60	3.69
0.9:0.1	7.32	7.63	6.37	3.51	1.82	4.74	7.20	2.08	4.33
Grouped by concept drift speed									
Abr.	6.47	6.83	5.53	3.47	2.60	5.32	7.31	4.09	3.39
Grad.	6.69	6.94	6.16	3.67	2.39	4.77	7.50	2.96	3.93
Real	5.90	5.78	6.57	4.73	4.47	2.22	8.36	2.09	4.88
Grouped by streams									
Sine	7.94	8.18	6.39	3.17	2.24	3.31	6.43	3.77	3.56
Agr.	5.70	5.64	5.35	4.20	2.86	6.59	8.87	2.39	3.41
SEA	7.78	7.78	6.82	2.91	2.31	4.95	5.95	1.88	4.62
STA.	5.79	7.16	5.33	3.37	2.21	3.78	6.90	7.19	3.28
Airlines	7.77	6.13	8.73	5.00	1.00	2.00	7.37	3.57	3.43
NOAA	5.03	5.03	7.93	4.87	6.70	1.90	9.00	1.33	3.20
Covtype	4.26	3.96	6.44	5.54	6.69	1.00	8.45	2.63	6.17
INSECTS	7.19	7.39	6.30	4.73	4.73	2.15	9.00	1.96	4.70
Luxembourg	7.50	7.50	3.50	3.00	4.90	1.00	3.80	4.20	8.67
Ozone	7.50	7.50	3.87	3.10	4.63	1.00	9.00	2.10	5.77
PAKDD-2009	6.50	6.50	8.00	4.90	4.10	1.00	9.00	2.17	2.83
Amazon	6.17	5.23	7.70	3.17	2.40	7.47	9.00	1.00	3.90
Twitter	7.53	7.33	7.70	4.43	2.80	1.97	9.00	1.33	4.13
All	6.68	6.89	6.00	3.84	2.29	4.40	7.32	4.92	3.79

- The p-values of Friedman tests are all ≤2.2E-16.
- Highlighted ranks denote significant superior performance.
- "Abr": Aburpt; "Grad.": Gradual; "Agr.": Agrawal; "STA.": STAGGER; "Real": Real-world

TABLE VIII
MEDIAN OF PREDICTIVE PERFORMANCE OF APPROACHES IN TIME-DECAY G-MEAN, CLASS 0, AND CLASS 1 RECALLS

Groups	GH-VFDT$_d$	HD-VFDT$_d$	Oza-Bag$_d$	OOB$_d$	UOB$_d$	CSARF	VFC-SMOTE	SMOTE-OB	CDCMS.CIL$_{OS}^{GMean}$
Time-Decay G-Mean									
Grouped by imbalance ratio (Artificial data streams)									
0.5:0.5	85.87%	85.86%	86.46%	83.36%	84.48%	82.52%	79.12%	84.03%	85.99%
0.7:0.3	81.19%	81.22%	79.71%	83.1%	83.96%	80.58%	78.06%	83.63%	84.56%
0.8:0.2	74.21%	73.09%	74.19%	81.08%	81.44%	78.84%	73.81%	82.66%	80.55%
0.9:0.1	52.58%	52.14%	69.06%	74.59%	79.26%	74.61%	59.62%	81.19%	72.78%
Grouped by concept drift speed									
Abr.	79.78%	80.0%	80.54%	81.28%	80.81%	79.84%	76.35%	83.31%	83.34%
Grad.	77.7%	77.85%	79.56%	80.97%	80.54%	79.75%	75.38%	82.5%	82.53%
Real	5.74	5.79	6.10	4.39	5.51	1.78	8.54	2.24	4.92
Grouped by streams									
Sine	83.14%	83.16%	85.02%	90.91%	90.03%	92.86%	86.05%	89.51%	90.05%
Agr.	73.7%	73.43%	75.4%	76.99%	76.71%	72.51%	47.87%	79.26%	78.9%
SEA	64.26%	64.26%	60.96%	79.93%	79.6%	79.1%	76.25%	80.14%	79.81%
STA.	98.66%	98.41%	98.93%	99.08%	99.03%	98.56%	98.8%	97.28%	99.03%
Airlines	6.00	7.10	7.93	4.70	4.30	2.30	8.97	2.70	1.00
NOAA	6.07	6.07	7.93	3.97	5.73	1.00	9.00	2.57	2.67
Covtype	3.89	3.88	6.47	5.36	6.97	1.00	8.40	2.57	6.47
INSECTS	7.31	7.20	5.12	3.82	4.79	2.32	9.00	1.39	4.04
Luxembourg	7.47	7.47	2.80	2.37	6.03	1.00	4.37	4.50	9.00
Ozone	7.50	7.50	3.57	2.93	5.73	1.00	9.00	2.93	4.83
PAKDD-2009	6.50	6.50	8.00	4.80	4.20	1.00	9.00	2.37	2.63
Amazon	6.07	4.93	7.20	2.90	3.07	7.33	9.00	1.00	3.50
Twitter	5.57	7.40	7.60	5.10	2.77	1.67	9.00	1.60	4.30
All	76.57%	76.74%	77.94%	80.47%	80.0%	78.9%	70.83%	80.88%	81.02%
Time-Decay Class 0 (Majority Class) Recall									
Grouped by imbalance ratio (Artificial data streams)									
0.5:0.5	86.0%	85.8%	87.4%	87.%	87.0%	86.1%	75.2%	85.8%	87.0%
0.7:0.3	93.4%	93.9%	95.9%	90.2%	85.5%	92.5%	92.6%	88.0%	91.2%
0.8:0.2	95.5%	96.2%	97.7%	90.6%	85.3%	93.9%	97.4%	87.7%	92.4%
0.9:0.1	98.5%	98.6%	99.6%	93.9%	84.2%	96.2%	99.3%	88.5%	94.1%
Grouped by concept drift speed									
Abr.	95.7%	96.2%	97.7%	91.8%	86.1%	94.5%	94.4%	87.1%	92.1%
Grad.	95.5%	96.1%	97.6%	91.4%	85.2%	93.7%	94.3%	87.7%	91.6%
Real	4.39	4.67	3.73	4.89	7.03	5.79	2.53	7.00	4.97
Grouped by streams									
Sine	97.4%	97.4%	98.4%	94.9%	89.6%	99.1%	94.1%	89.6%	95.5%
Agr.	93.6%	94.5%	96.7%	89.3%	82.2%	92.1%	94.9%	87.0%	90.9%
SEA	89.3%	89.3%	90.1%	75.8%	72.4%	76.8%	78.2%	72.5%	77.7%
STA.	99.7%	99.7%	99.8%	99.6%	99.1%	99.5%	99.8%	97.9%	99.7%
Airlines	1.67	3.27	1.40	4.60	9.00	7.27	5.13	7.13	5.53
NOAA	7.97	7.97	2.00	4.87	4.33	3.63	1.00	8.03	5.20
Covtype	5.13	5.41	5.90	4.14	6.69	1.00	8.45	2.57	6.47
INSECTS	3.77	3.76	2.16	5.22	8.67	6.15	1.00	7.97	6.31
Luxembourg	7.47	7.47	2.68	2.45	4.53	1.03	5.62	4.78	8.97
Ozone	2.50	2.50	4.97	9.00	4.03	7.00	1.00	8.00	6.00
PAKDD-2009	3.50	3.50	2.00	5.10	5.90	7.00	1.00	9.00	8.00
Amazon	1.63	4.57	3.43	6.13	7.80	3.40	2.03	9.00	7.00
Twitter	4.00	2.77	2.23	5.60	7.83	9.00	1.00	7.13	5.43
All	94.68%	95.18%	97.14%	90.54%	85.06%	93.69%	94.25%	87.21%	91.05%
Time-Decay Class 1 (Minority Class) Recall									
Grouped by imbalance ratio (Artificial data streams)									
0.5:0.5	88.1%	88.2%	90.1%	89.9%	88.9%	88.7%	86.3%	88.5%	89.9%
0.7:0.3	71.2%	71.1%	67.3%	86.0%	88.6%	83.8%	78.5%	89.0%	84.9%
0.8:0.2	58.8%	57.5%	57.1%	82.7%	87.3%	79.0%	69.9%	88.4%	80.3%
0.9:0.1	30.6%	30.3%	49.0%	67.3%	83.5%	69.3%	46.4%	86.9%	63.5%
Grouped by concept drift speed									
Abr.	73.4%	73.6%	74.7%	86.1%	88.0%	81.9%	73.4%	88.4%	83.9%
Grad.	68.9%	68.3%	70.1%	83.2%	87.7%	78.4%	70.8%	88.2%	81.9%
Real	5.90	5.78	6.57	4.73	4.47	2.22%	8.36	2.09	4.88
Grouped by streams									
Sine	71.2%	71.3%	74.0%	88.7%	90.8%	87.4%	79.9%	89.4%	88.0%
Agr.	61.0%	60.4%	61.3%	69.1%	72.5%	58.9%	27.0%	74.1%	69.3%
SEA	52.3%	52.3%	49.6%	85.7%	88.0%	82.0%	75.7%	89.7%	82.9%
STA.	98.5%	98.4%	98.7%	99.0%	99.1%	98.5%	98.5%	96.7%	98.8%
Airlines	7.77	6.13	8.73	5.00	1.00	2.00	7.37	3.57	3.43
NOAA	5.03	5.03	7.93	4.87	6.70	1.90	9.00	1.33	3.20
Covtype	4.26	3.96	6.44	5.54	6.69	1.00	8.45	2.63	6.17
INSECTS	7.19	7.39	6.30	4.73	4.73	2.15	9.00	1.96	4.70
Luxembourg	7.50	7.50	3.50	3.00	4.90	1.00	3.80	4.20	8.67
Ozone	7.50	7.50	3.87	3.10	4.63	1.00	9.00	2.10	5.77
PAKDD-2009	6.50	6.50	8.00	4.90	4.10	1.00	9.00	2.17	2.83
Amazon	6.17	5.23	7.70	3.17	2.40	7.47	9.00	1.00	3.90
Twitter	7.53	7.33	7.70	4.43	2.80	1.97	9.00	1.33	4.13
All	66.73%	66.2%	65.95%	80.71%	85.92%	75.44%	64.75%	87.31%	80.25%

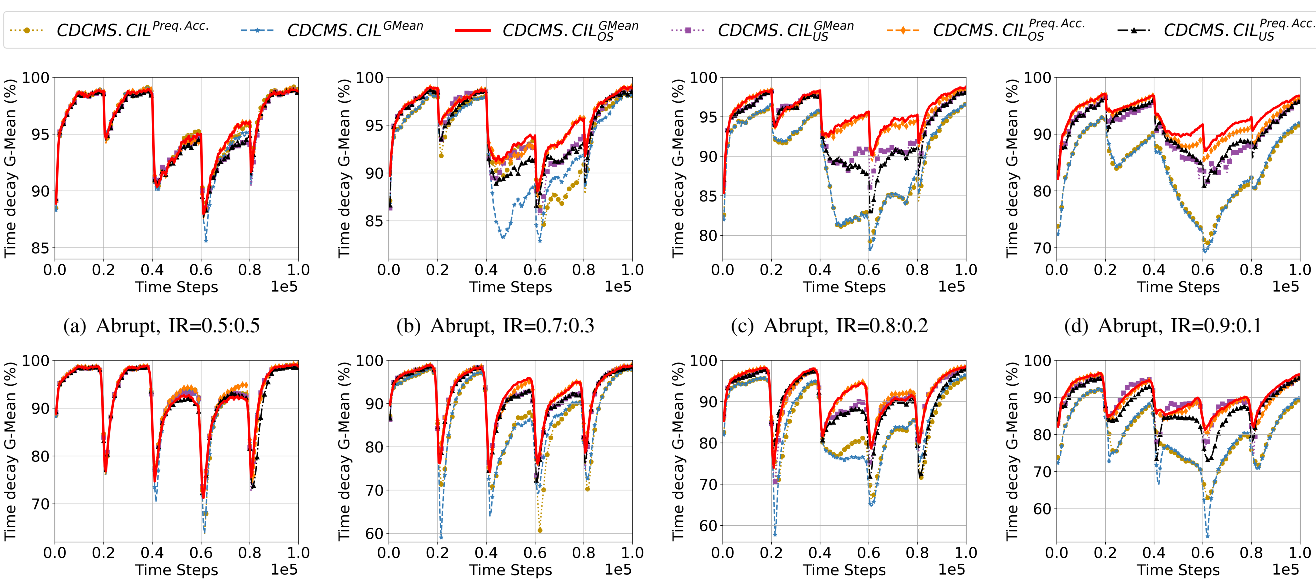
- The p-values of Friedman tests are all ≤2.2E-16.
- Highlighted ranks denote significant superior performance.
- "Abr": Aburpt; "Grad.": Gradual; "Agr.": Agrawal; "STA.": STAGGER; "Real": Real-world



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

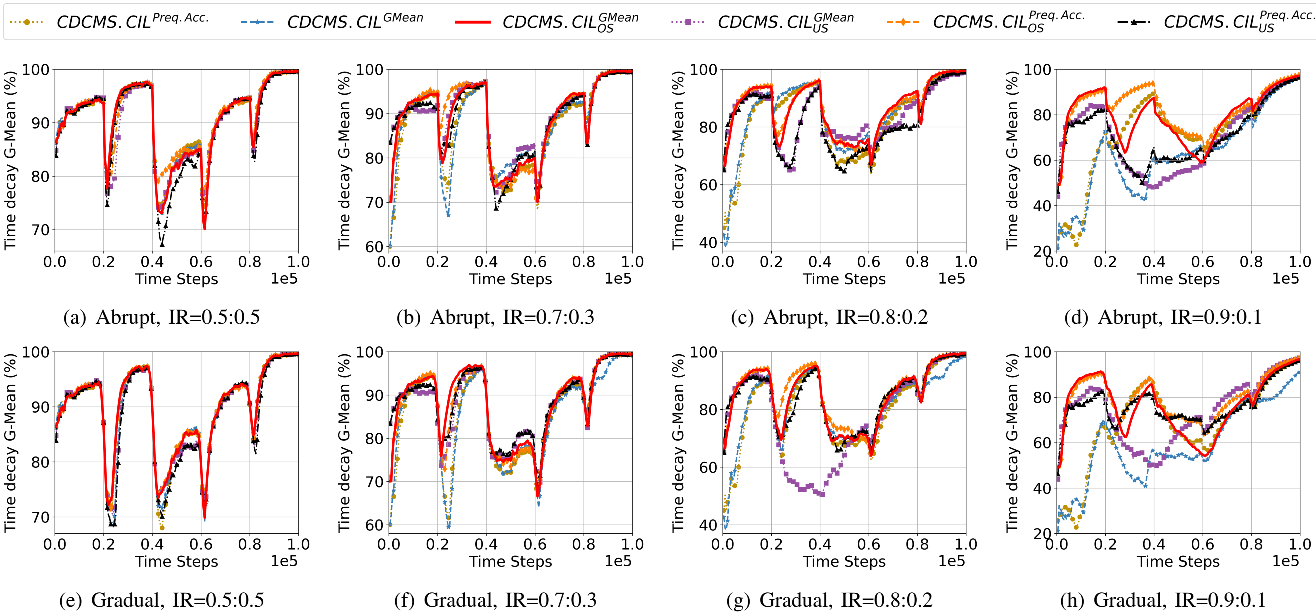
Fig. 1. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Sine1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

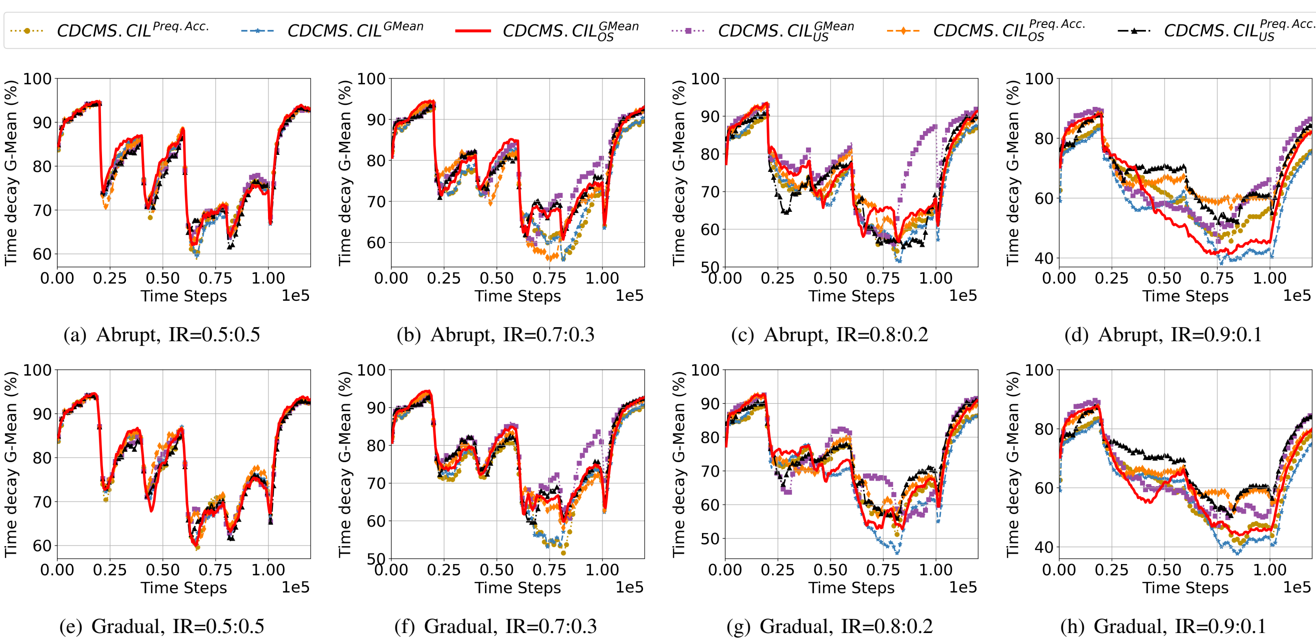
Fig. 2. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Sine2 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

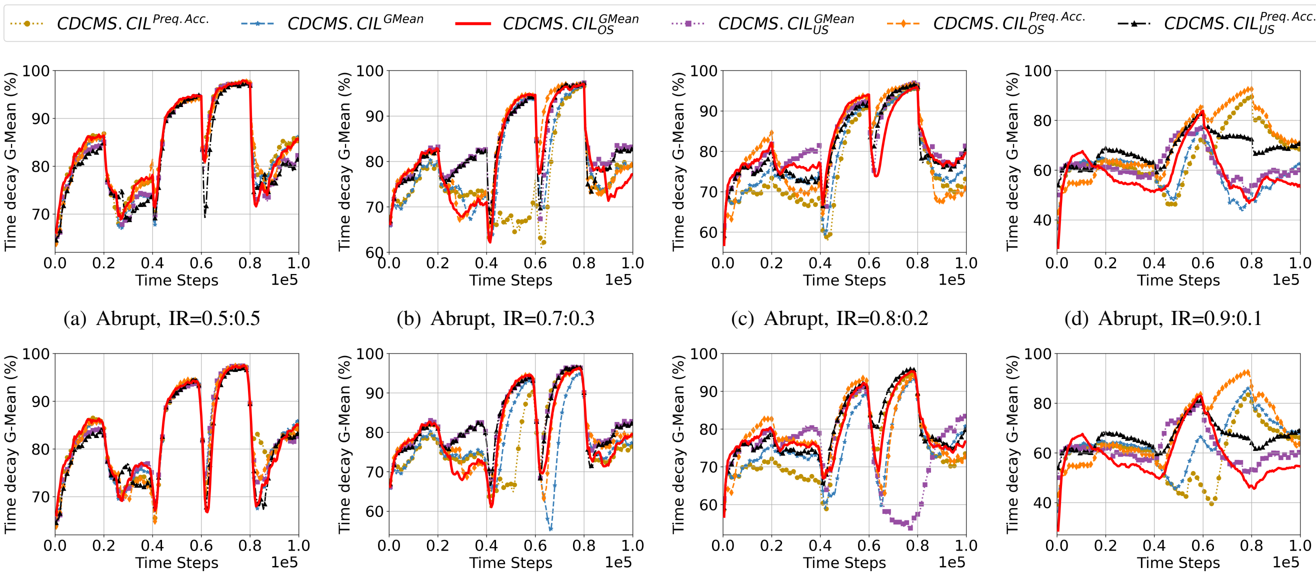
Fig. 3. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

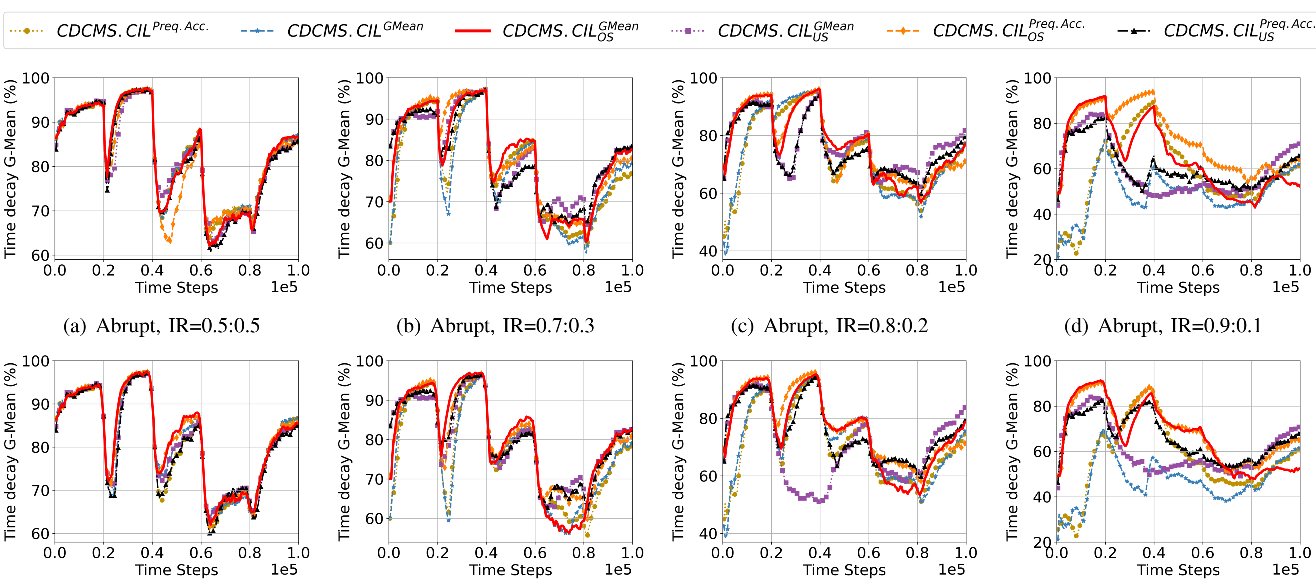
Fig. 4. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal2 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

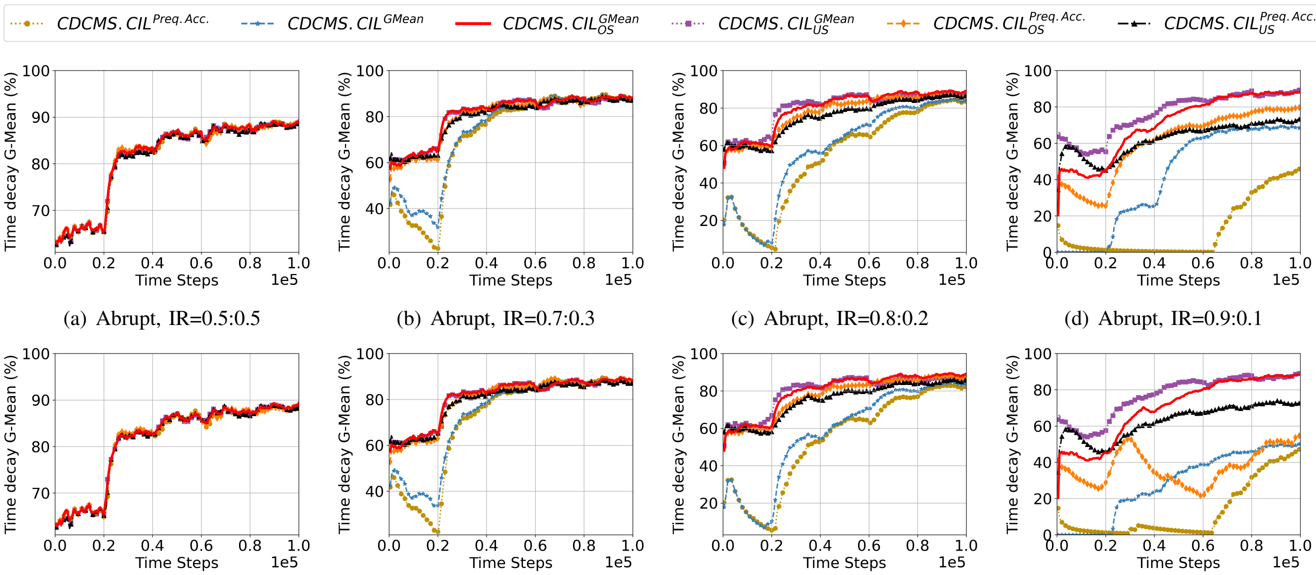
Fig. 5. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal3 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

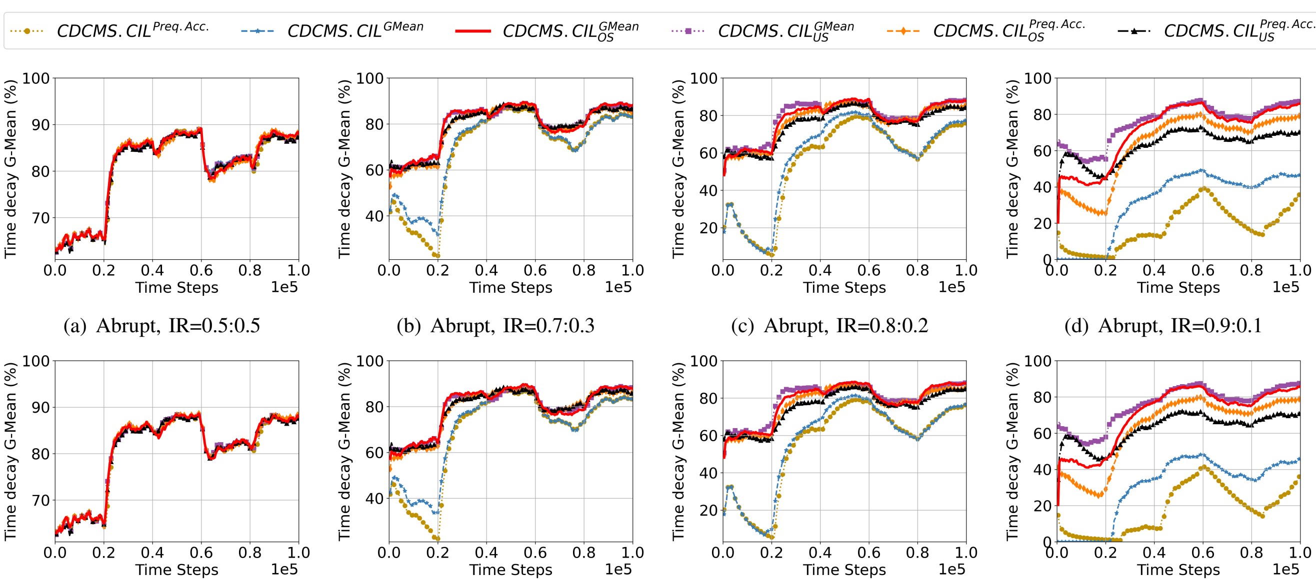
Fig. 6. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal4 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

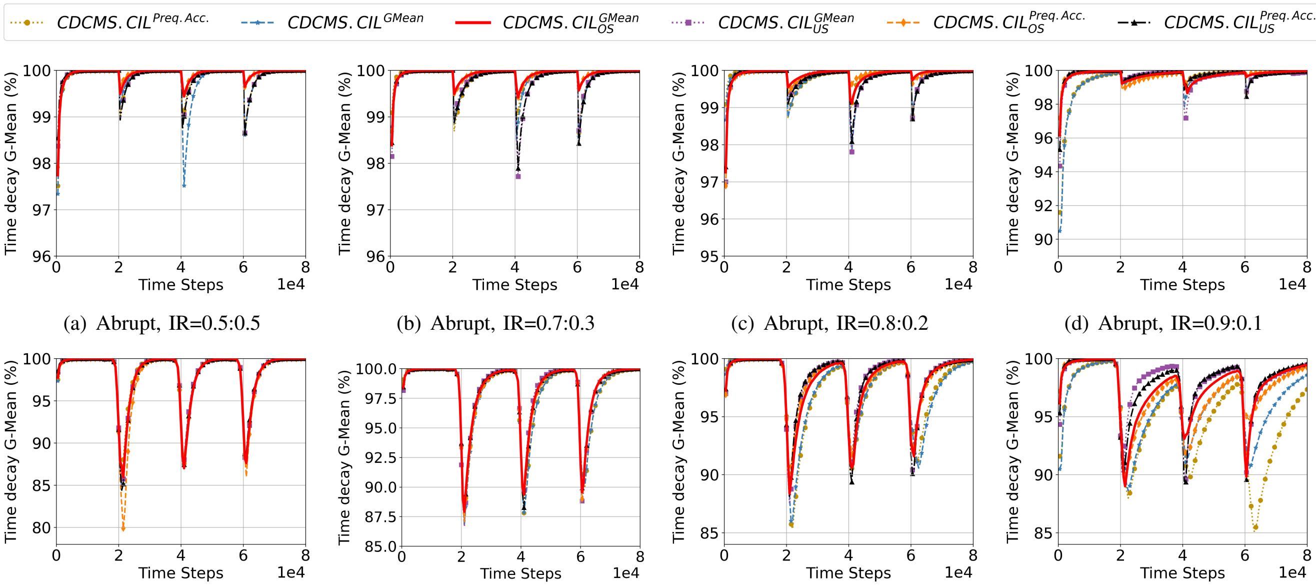
Fig. 7. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on SEA1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

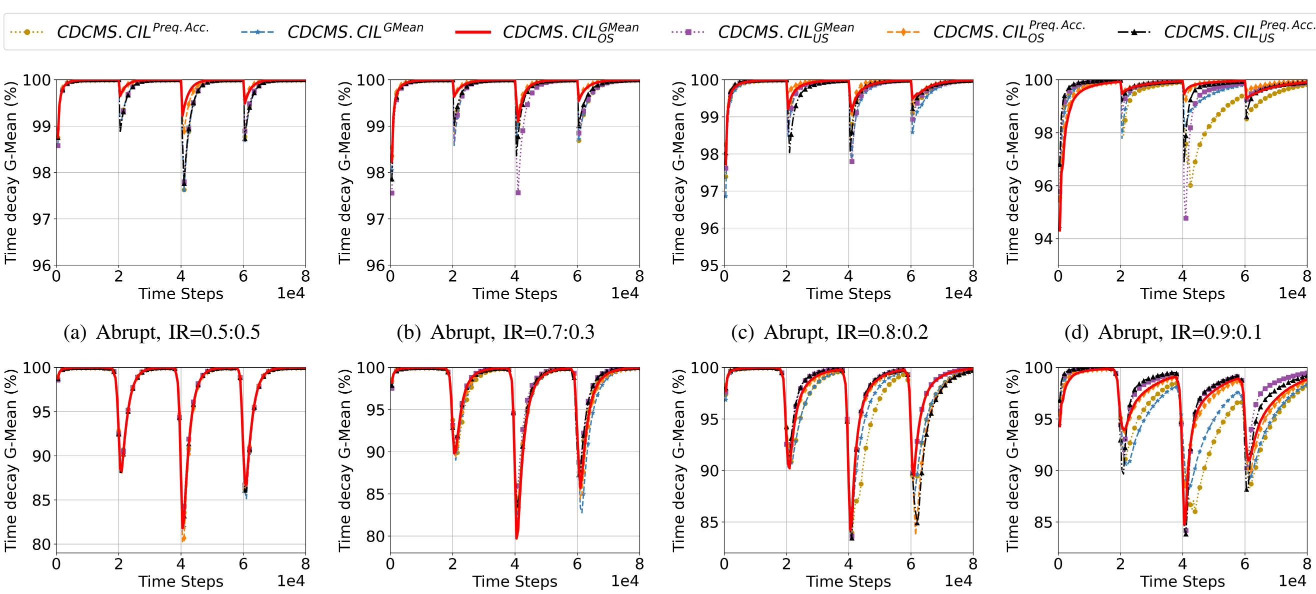
Fig. 8. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on SEA2 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

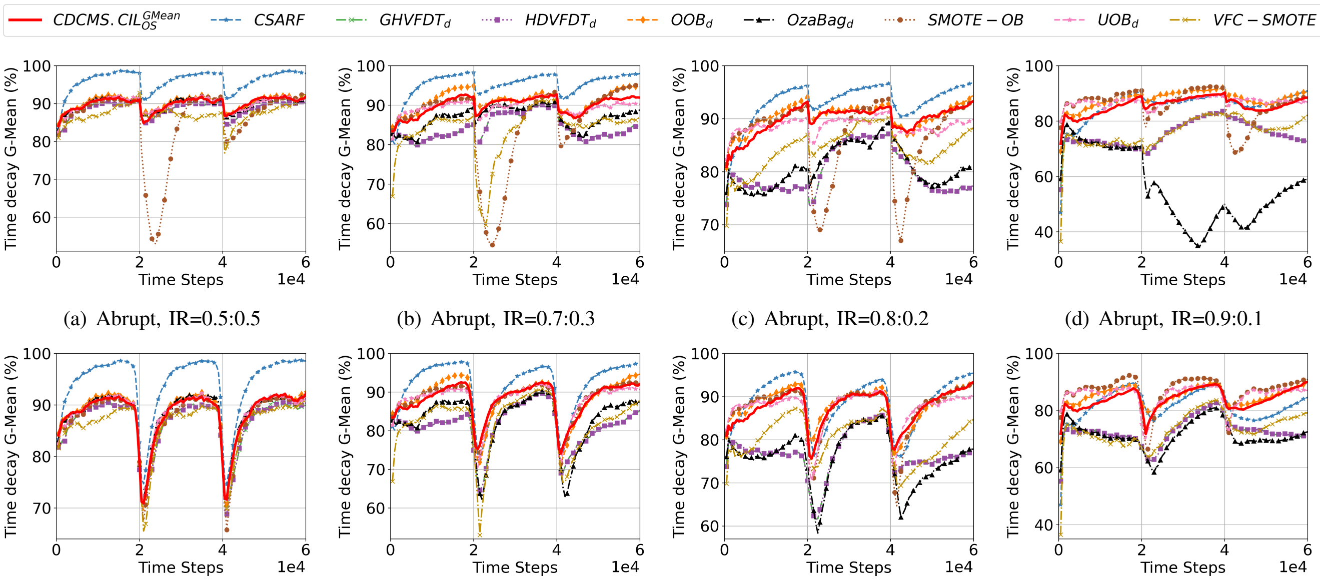
Fig. 9. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on STAGGER1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

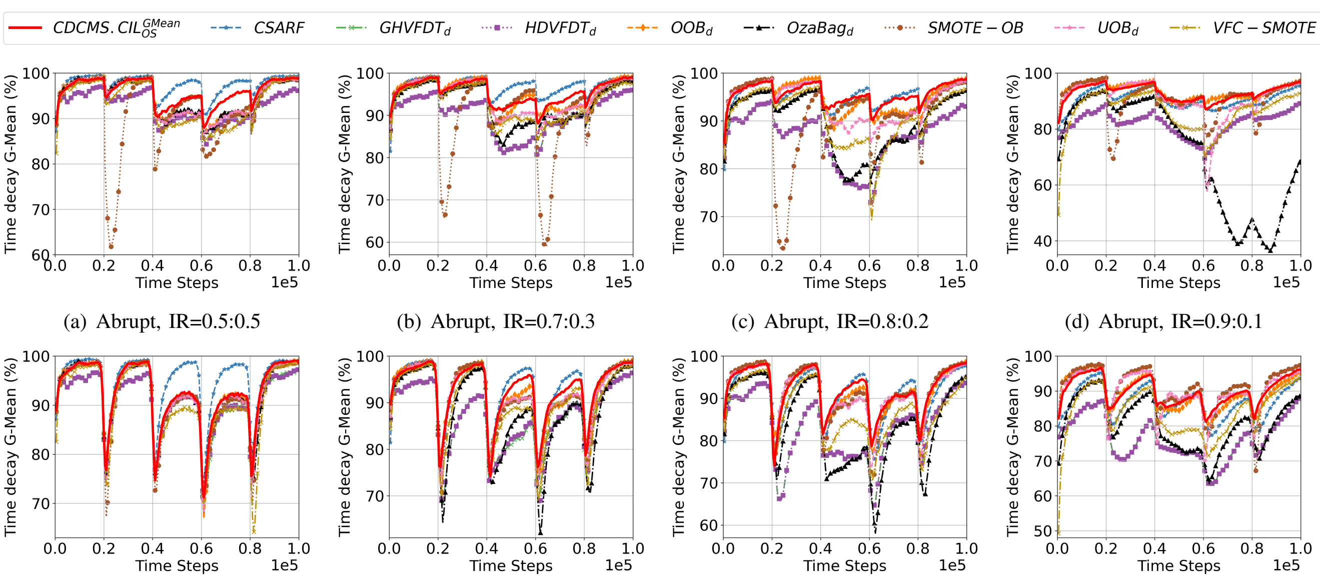
(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

Fig. 10. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on STAGGER2 (Time-decay G-Mean)



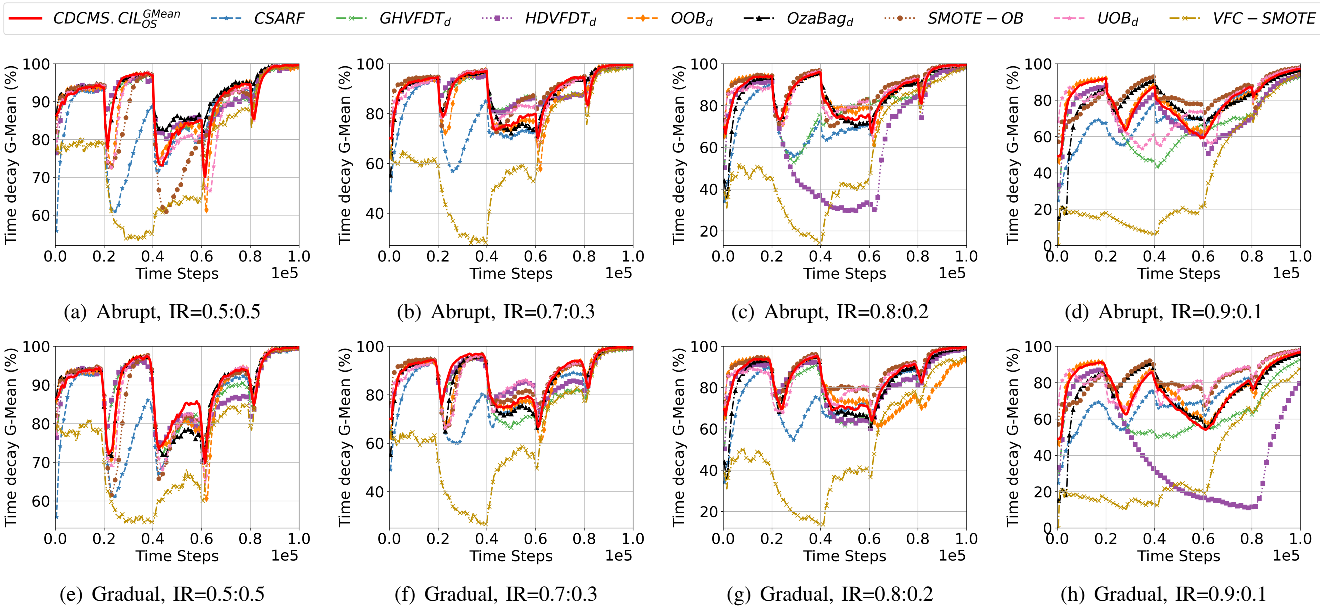
(a) Abrupt, IR=0.5:0.5
(b) Abrupt, IR=0.7:0.3
(c) Abrupt, IR=0.8:0.2
(d) Abrupt, IR=0.9:0.1
(e) Gradual, IR=0.5:0.5
(f) Gradual, IR=0.7:0.3
(g) Gradual, IR=0.8:0.2
(h) Gradual, IR=0.9:0.1

Fig. 11. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Sine1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5
(b) Abrupt, IR=0.7:0.3
(c) Abrupt, IR=0.8:0.2
(d) Abrupt, IR=0.9:0.1
(e) Gradual, IR=0.5:0.5
(f) Gradual, IR=0.7:0.3
(g) Gradual, IR=0.8:0.2
(h) Gradual, IR=0.9:0.1

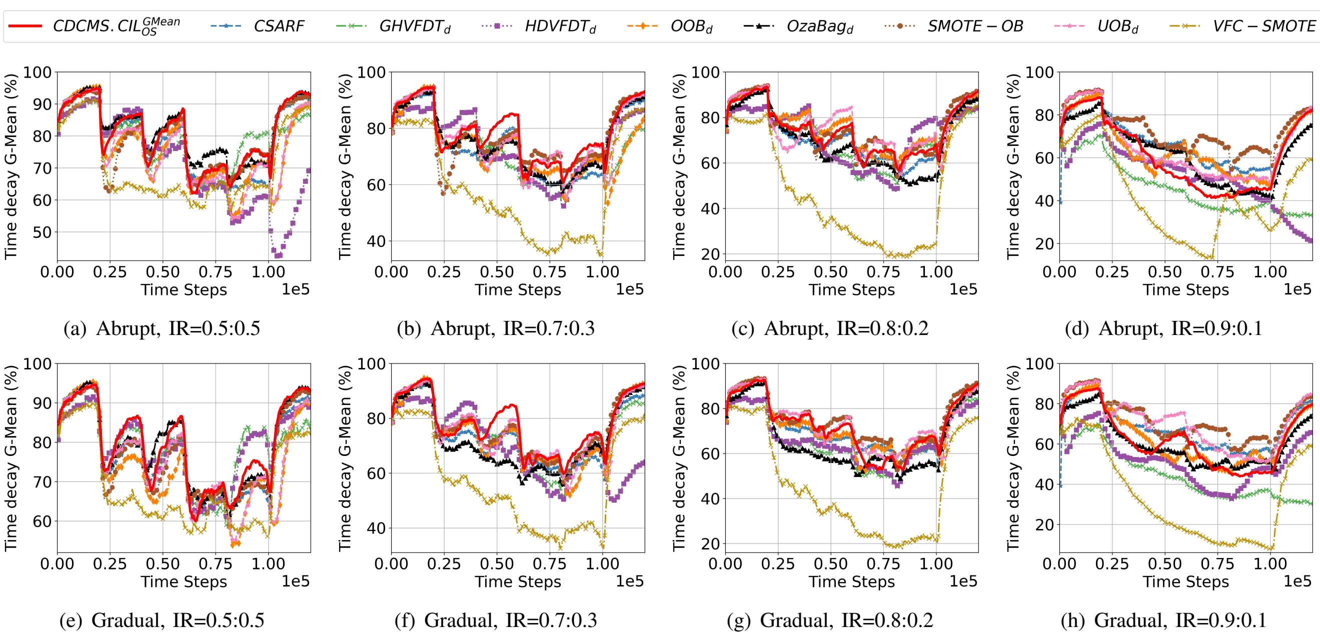
Fig. 12. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Sine2 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

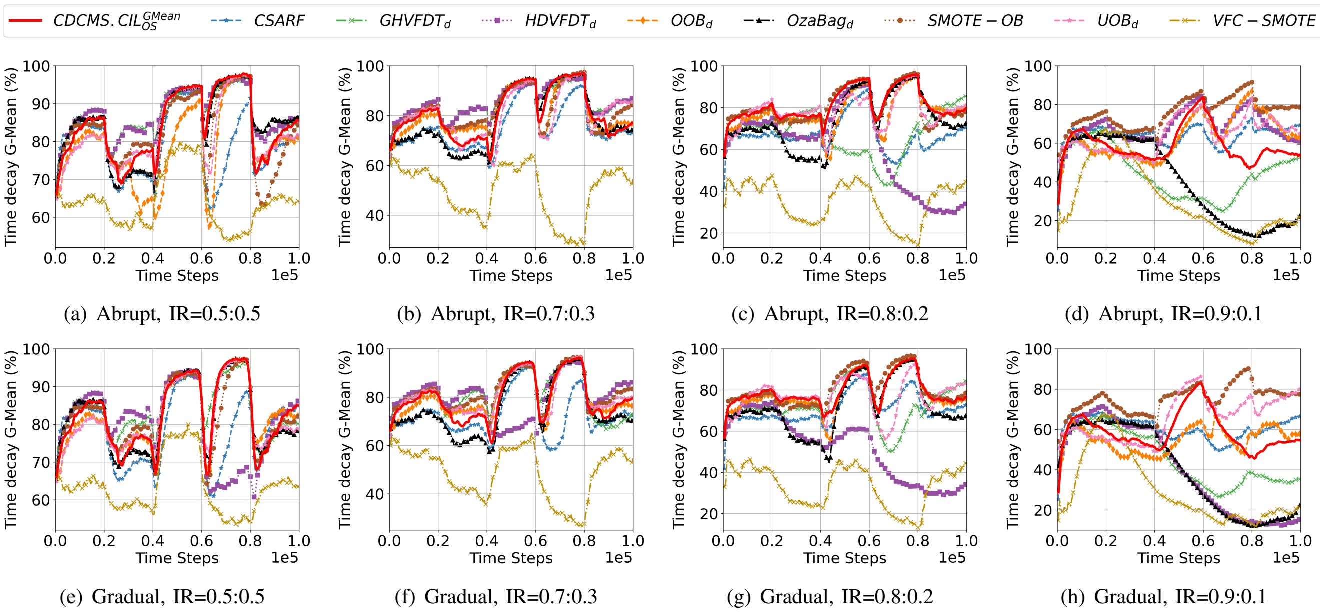
Fig. 13. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

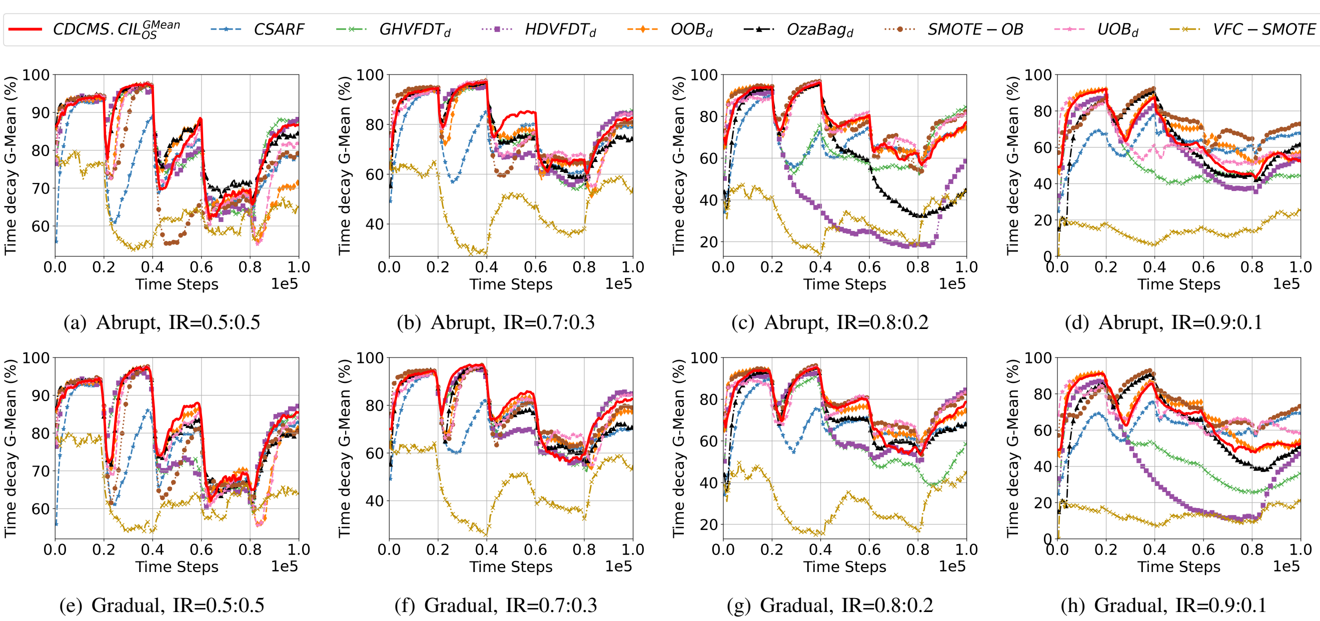
Fig. 14. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal2 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

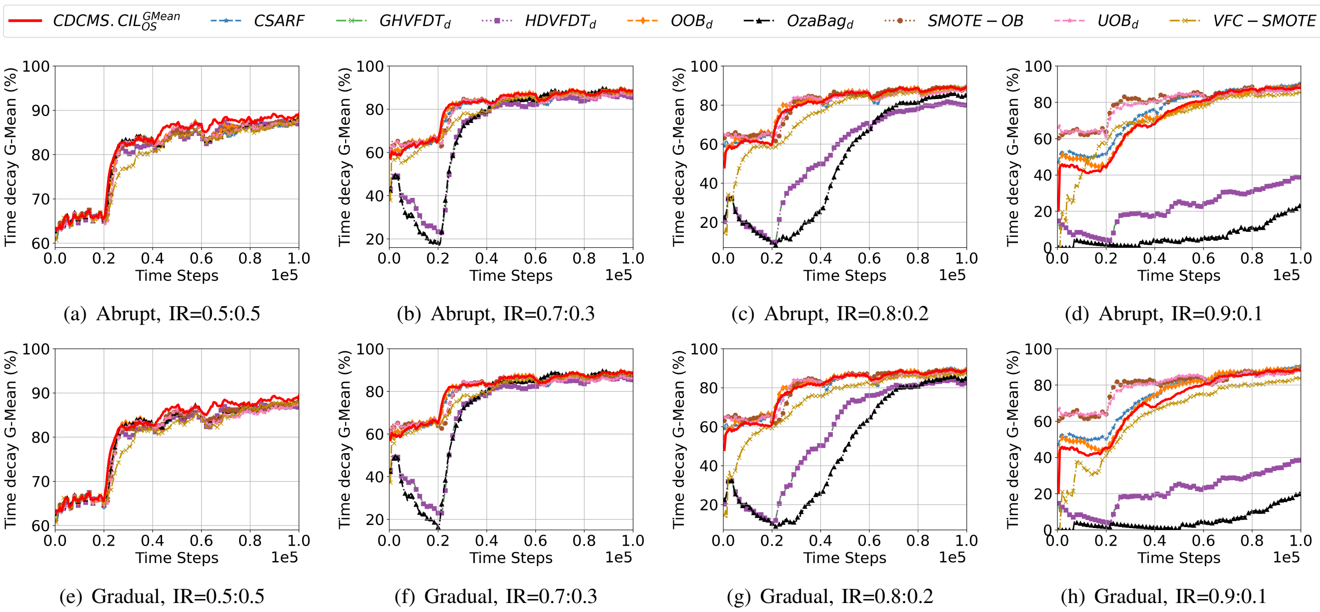
Fig. 15. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal3 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

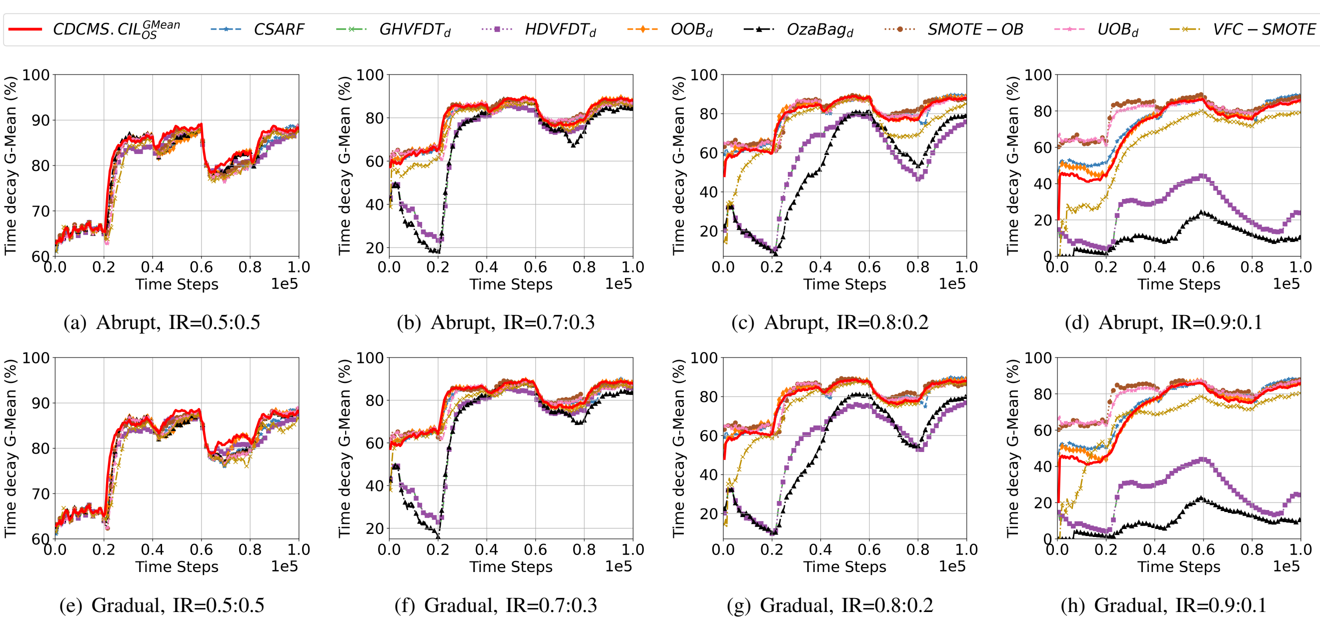
Fig. 16. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Agrawal4 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

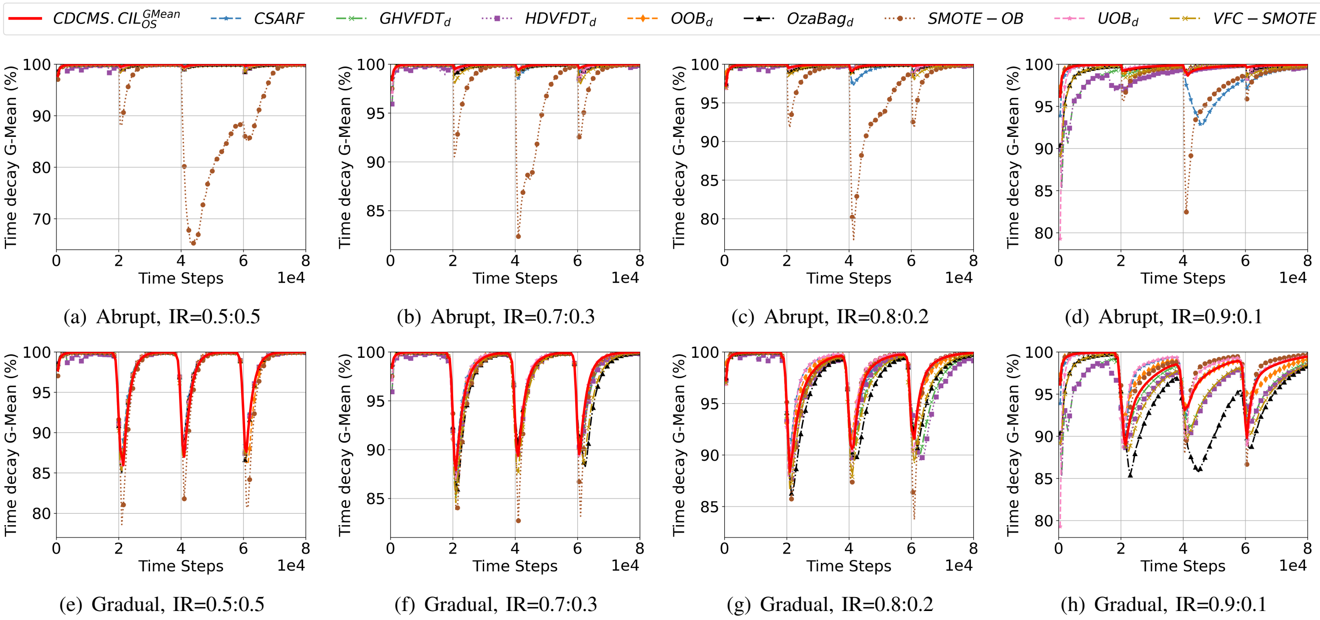
Fig. 17. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on SEA1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

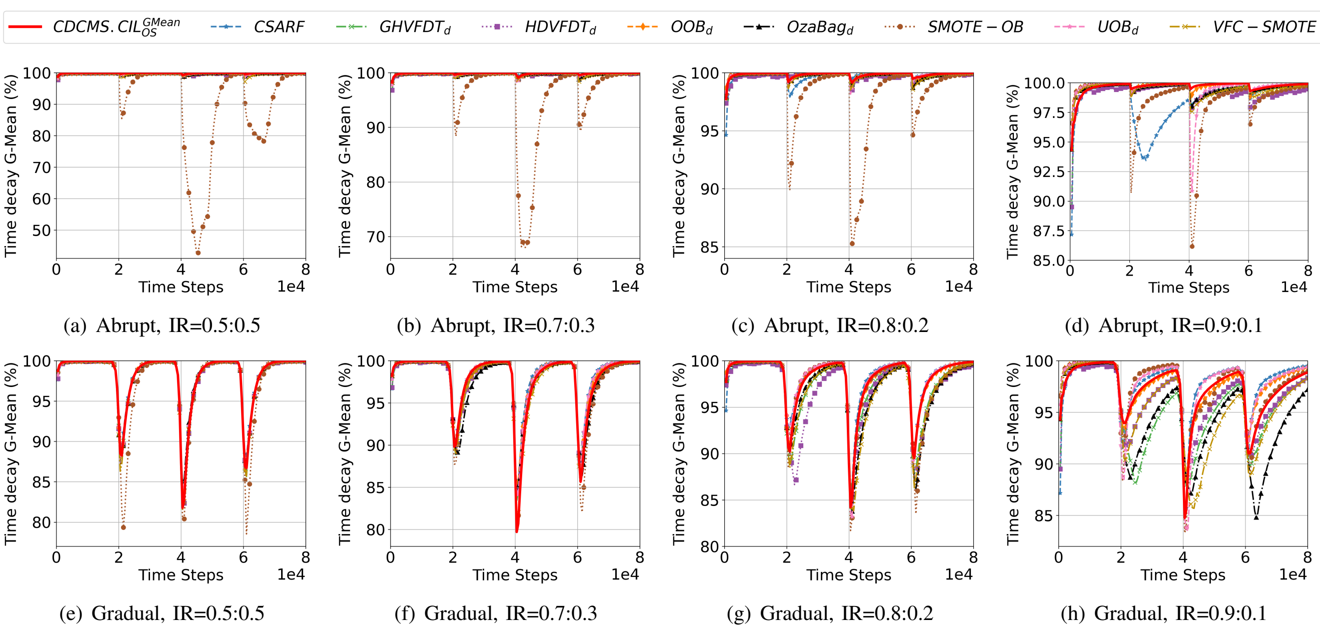
Fig. 18. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on SEA2 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

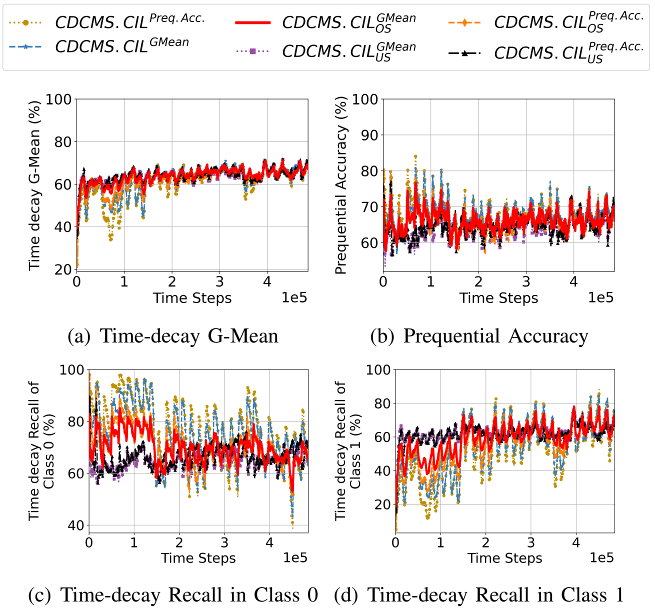
Fig. 19. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on STAGGER1 (Time-decay G-Mean)



(a) Abrupt, IR=0.5:0.5 (b) Abrupt, IR=0.7:0.3 (c) Abrupt, IR=0.8:0.2 (d) Abrupt, IR=0.9:0.1

(e) Gradual, IR=0.5:0.5 (f) Gradual, IR=0.7:0.3 (g) Gradual, IR=0.8:0.2 (h) Gradual, IR=0.9:0.1

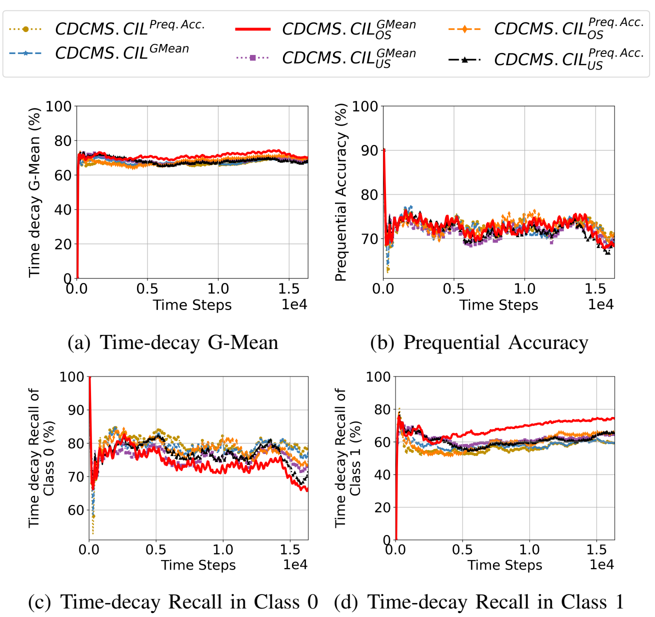
Fig. 20. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on STAGGER2 (Time-decay G-Mean)



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

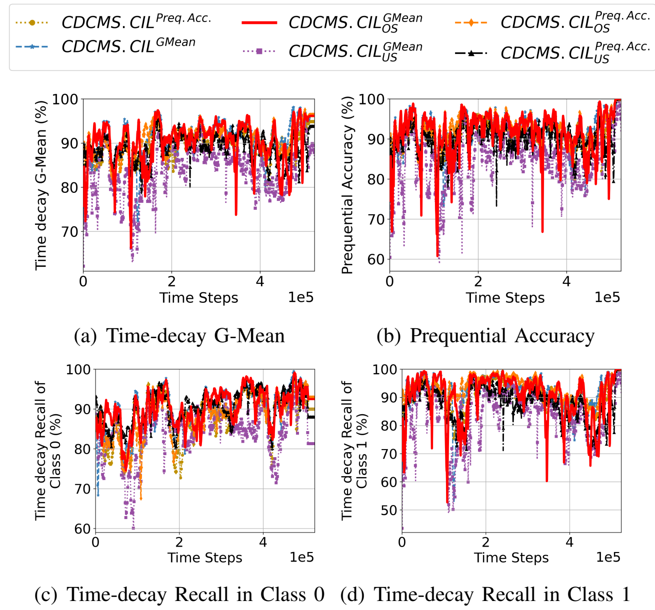
Fig. 21. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Airlines



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

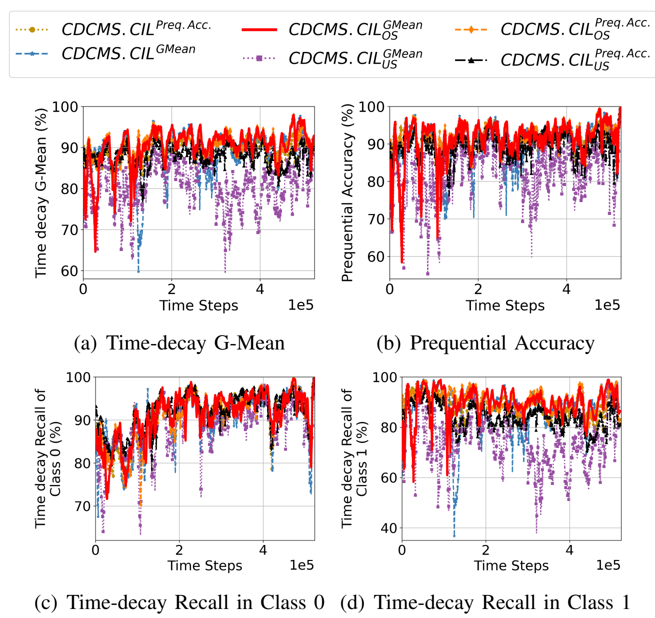
Fig. 22. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on NOAA



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 23. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Covtype$_{(c_1=1\text{-}6)}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 24. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Covtype$_{(c_1=1)}$

References

[1] G. U. Yule, "On the association of attributes in statistics," *Philosophical Transactions*, vol. A, no. 194, pp. 257–319, 1900.

[2] C. W. Chiu and L. L. Minku, "A Diversity Framework for Dealing With Multiple Types of Concept Drift Based on Clustering in the Model Space," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 3, pp. 1299–1309, 2022.

[3] A. Bifet, G. Holmes, R. Kirkby, and B. Pfahringer, "MOA: Massive Online Analysis," *Journal of Machine Learning Research*, vol. 11, no. 52, pp. 1601–1604, 2010. [Online]. Available: http://jmlr.org/papers/v11/bifet10a.html

[4] J. Gama, P. Medas, G. Castillo, and P. Rodrigues, "Learning with Drift Detection," vol. 8, 09 2004, pp. 286–295.

[5] R. Agrawal, T. Imielinski, and A. Swami, "Database mining: A performance perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 5, no. 6, pp. 914–925, 1993.

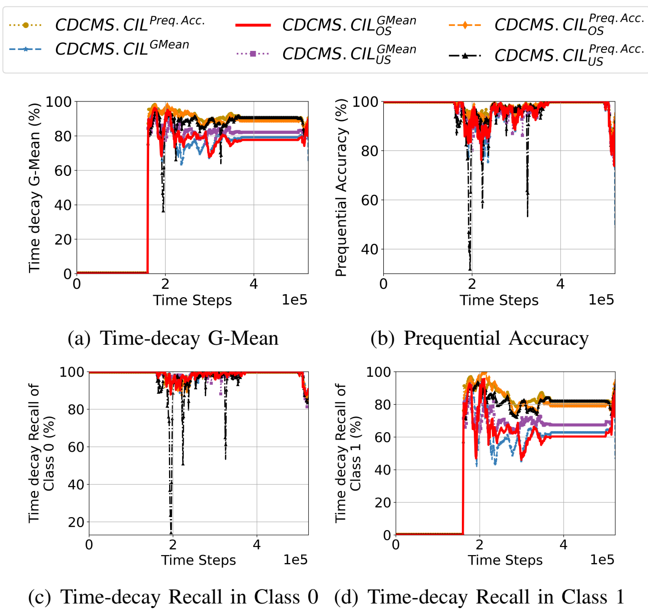
[6] W. N. Street and Y. Kim, "A Streaming Ensemble Algorithm (SEA) for Large-scale Classification," 2001, pp. 377–382.

[7] J. C. Schlimmer and R. H. Granger Jr., "Incremental Learning from Noisy Data," *Machine Learning*, vol. 1, pp. 317–354, 09 1986.

[8] V. M. A. Souza, D. M. dos Reis, A. G. Maletzke, and G. E. A. P. A. Batista, "Challenges in Benchmarking Stream Learning Algorithms with Real-world Data," *Data Mining and Knowledge Discovery*, vol. 34, pp. 1805–1858, 2020.

[9] E. Ikonomovska, J. a. Gama, and S. Džeroski, "Learning model trees from evolving data streams," *Data Mining and Knowledge Discovery*, vol. 23, pp. 128–168, 07 2011.

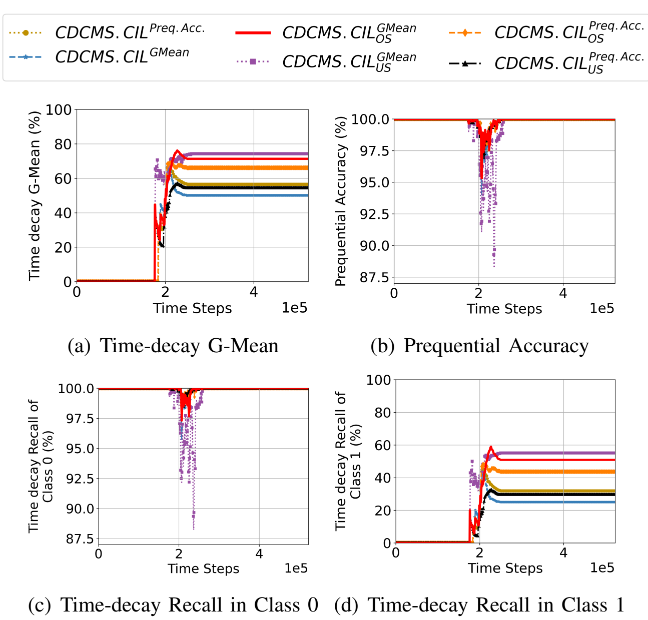
[10] R. Elwell and R. Polikar, "Incremental Learning of Concept Drift in



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

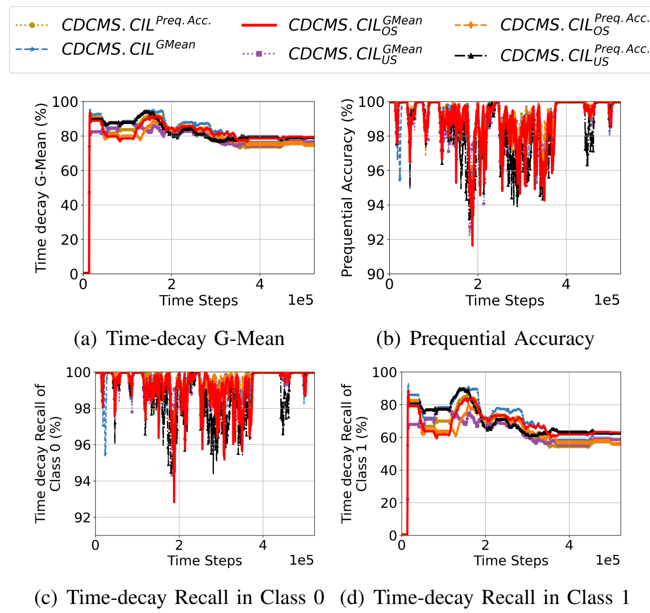
Fig. 25. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Covtype$_{(c_1=2)}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

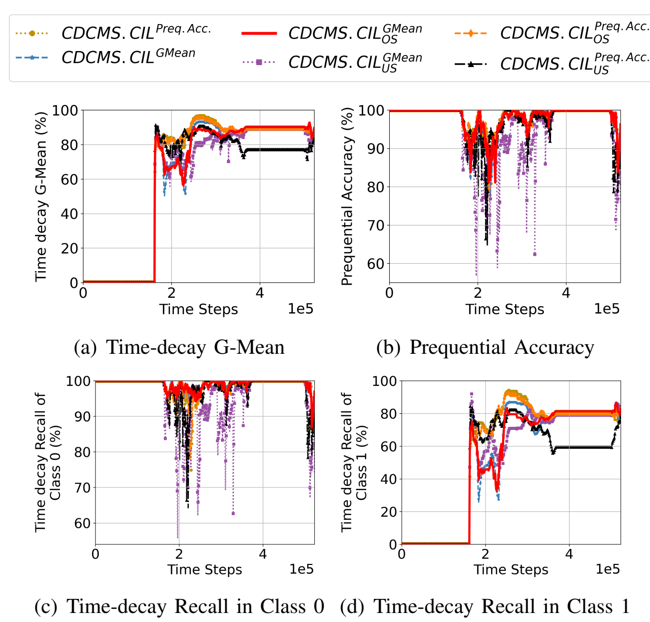
Fig. 26. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Covtype$_{(c_1=3)}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 27. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Covtype$_{(c_1=4)}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 28. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Covtype$_{(c_1=5)}$

Nonstationary Environments," *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1517–1531, 2011.

[11] J. A. Blackard and D. J. Dean, "Comparative Accuracies of Artificial Neural Networks and Discriminant Analysis in Predicting Forest Cover Types from Cartographic Variables," *Computers and Electronics in Agriculture*, vol. 24, pp. 131–151, 12 1999.

[12] I. Žliobaitė, "Combining similarity in time and space for training set formation under concept drift," *Intelligent Data Analysis*, vol. 15, pp. 589–611, 06 2011.

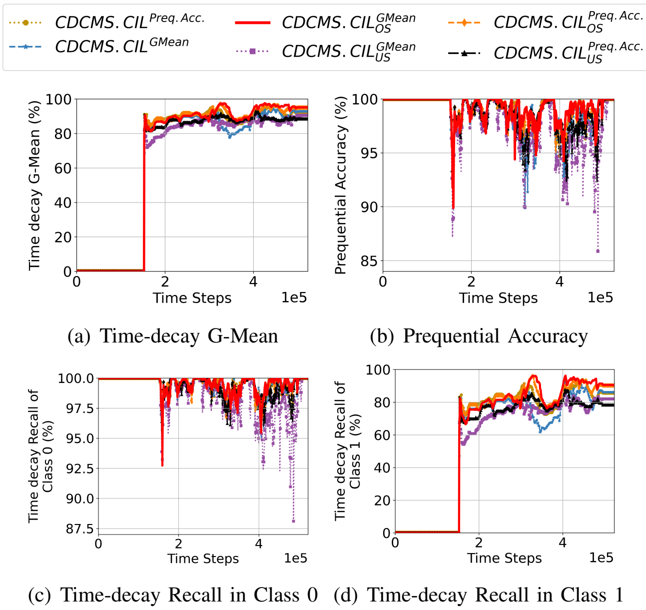
[13] K. Zhang, W. Fan, X. Yuan, I. Davidson, and X. Li, "Forecasting Skewed Biased Stochastic Ozone Days: Analyses and Solutions," vol. 14, 12 2006, pp. 753–764.

[14] T. Theeramunkong, B. Kijsirikul, N. Cercone, and T. B. Ho, Eds., *Advances in Knowledge Discovery and Data Mining, 13th Pacific-Asia Conference, PAKDD 2009, Bangkok, Thailand, April 27-30, 2009, Proceedings*, ser. Lecture Notes in Computer Science, vol. 5476. Springer, 2009. [Online]. Available: https://doi.org/10.1007/978-3-642-01307-2

[15] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, Bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification," in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, A. Zaenen and A. van den Bosch, Eds. Prague, Czech Republic: Association for Computational Linguistics, Jun. 2007, pp. 440–447. [Online]. Available: https://aclanthology.org/P07-1056

[16] P. Nakov, A. Ritter, S. Rosenthal, F. Sebastiani, and V. Stoyanov, "Semeval-2016 task 4: Sentiment analysis in twitter," *CoRR*, vol. abs/1912.01973, 2019. [Online]. Available: http://arxiv.org/abs/1912.01973

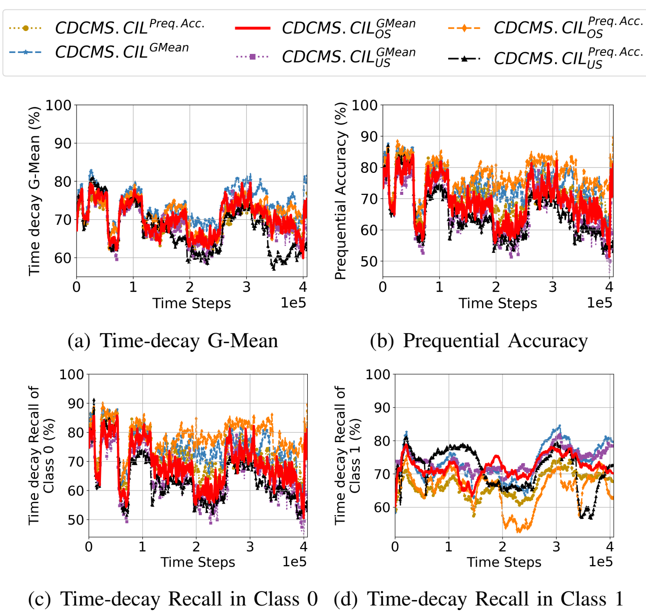
[17] P. Domingos and G. Hulten, "Mining High-Speed Data Streams," ser.



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 29. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on $\text{Covtype}_{(c_1=6)}$

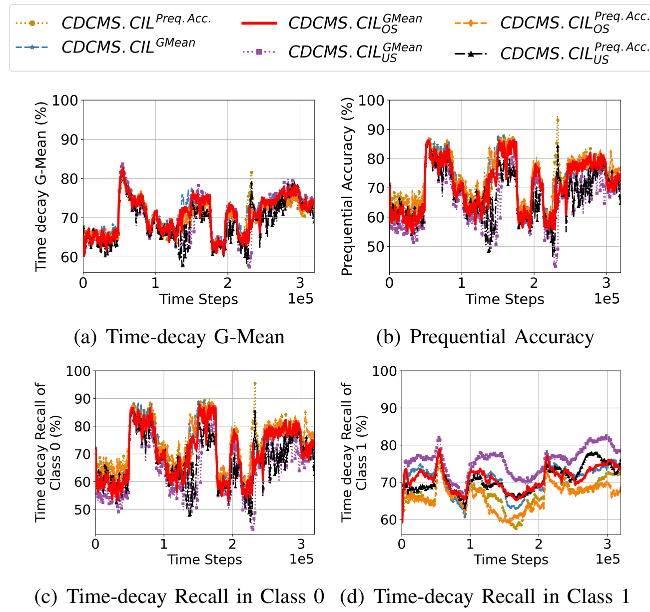


(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 30. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on $\text{INSECTS}^{\text{incremental}}$

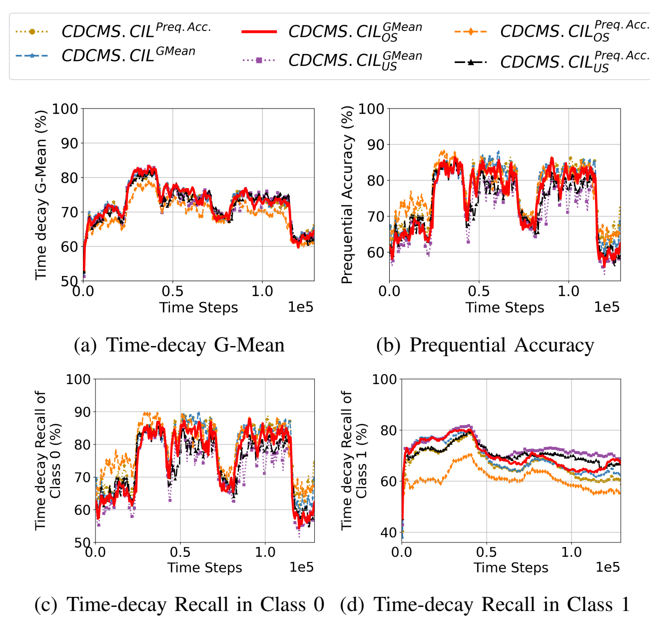
KDD '00. New York, NY, USA: Association for Computing Machinery, 2000, p. 71–80.



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

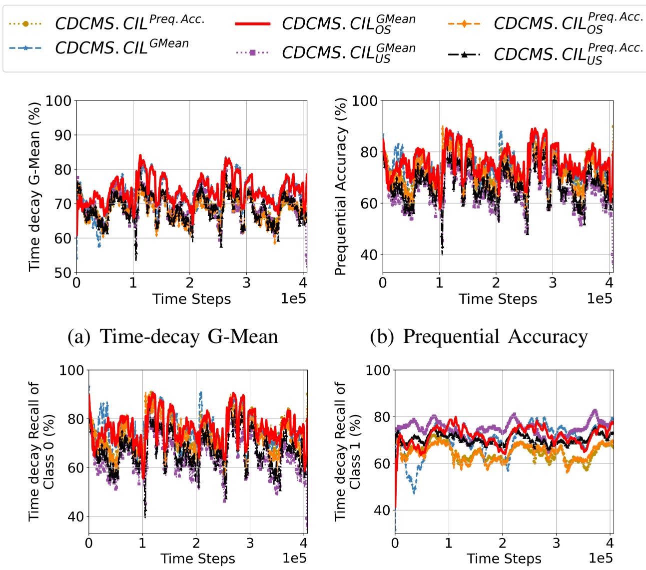
Fig. 31. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on $\text{INSECTS}_{\text{abrupt}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

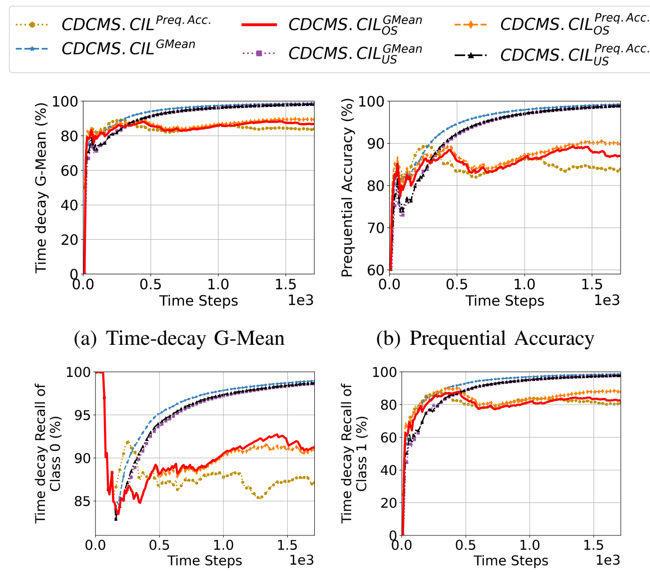
Fig. 32. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on $\text{INSECTS}^{\text{incremental}}_{\text{gradual}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

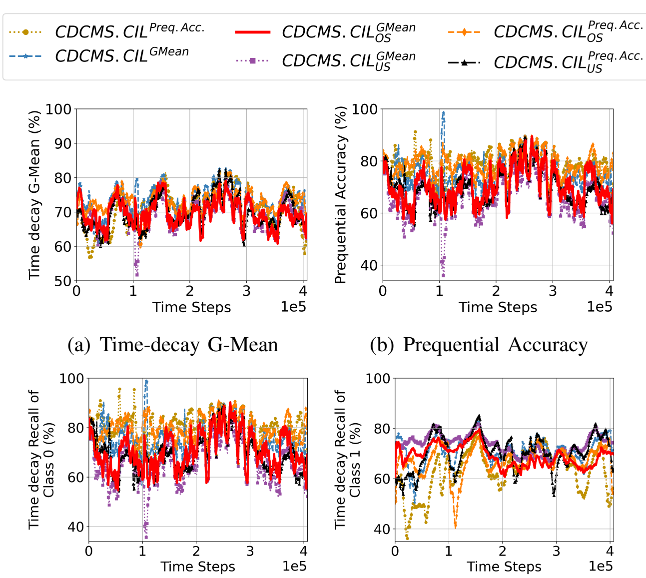
Fig. 33. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on $\text{INSECTS}^{\text{incremental}}_{\text{abrupt recurring}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

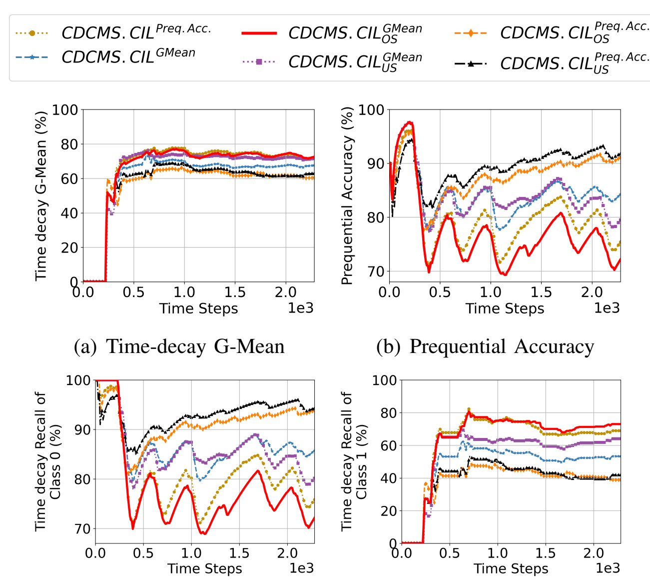
Fig. 35. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Luxembourg



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

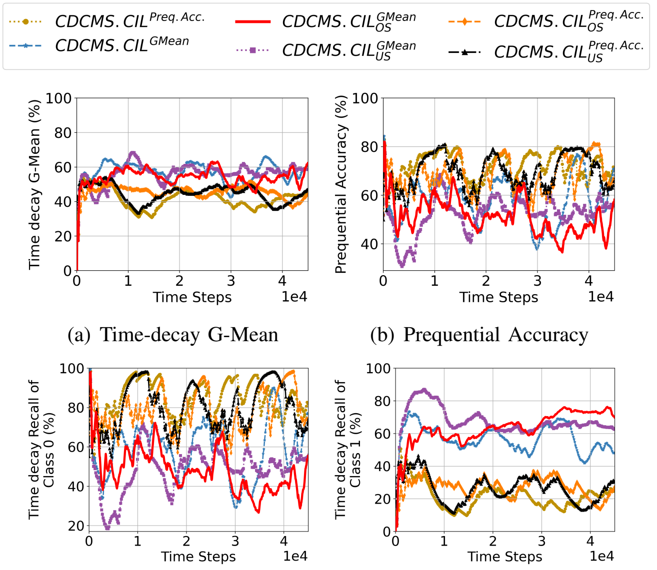
Fig. 34. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on $\text{INSECTS}^{\text{incremental}}_{\text{recurring}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

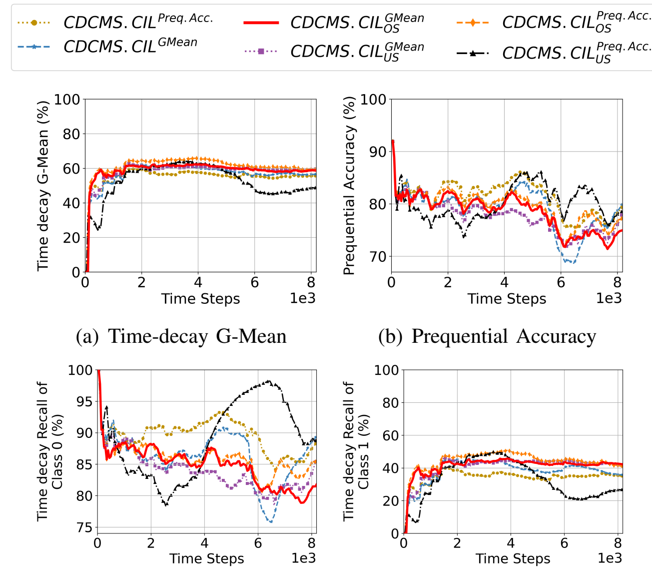
Fig. 36. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Ozone



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

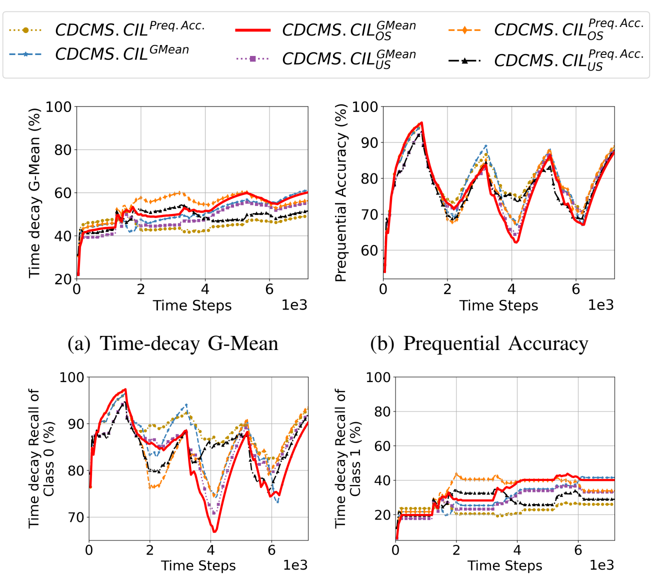
Fig. 37. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on PAKDD-2009



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

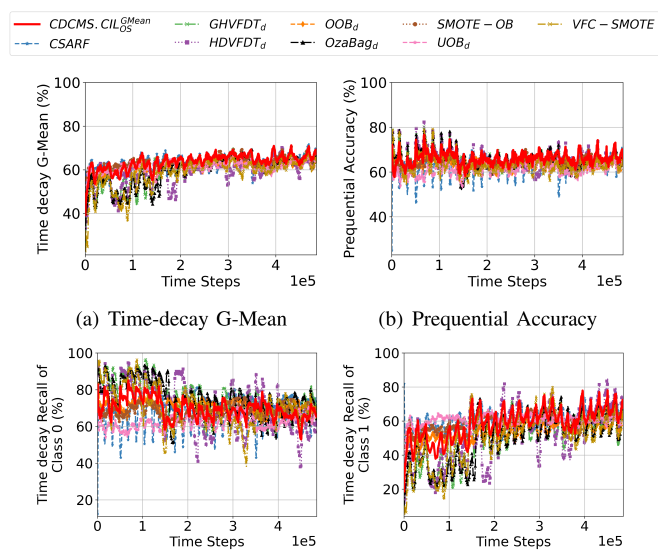
Fig. 39. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Twitter



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

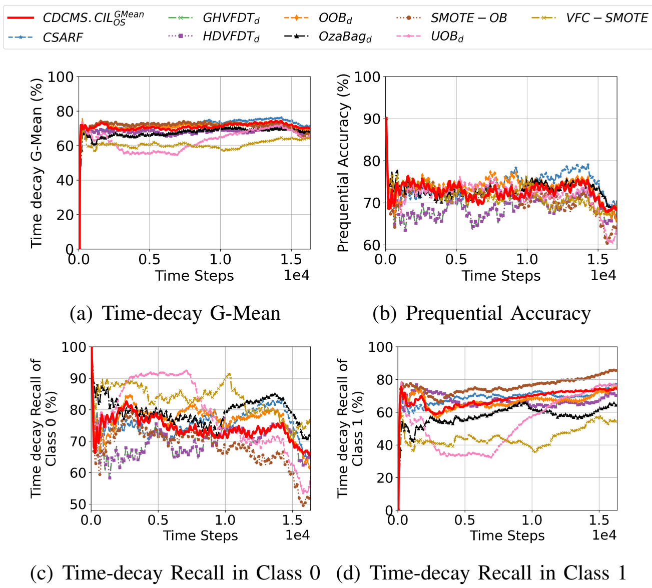
Fig. 38. Time-decay G-Mean of CDCMS.CIL with different weighting metrics and resampling strategies on Amazon



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

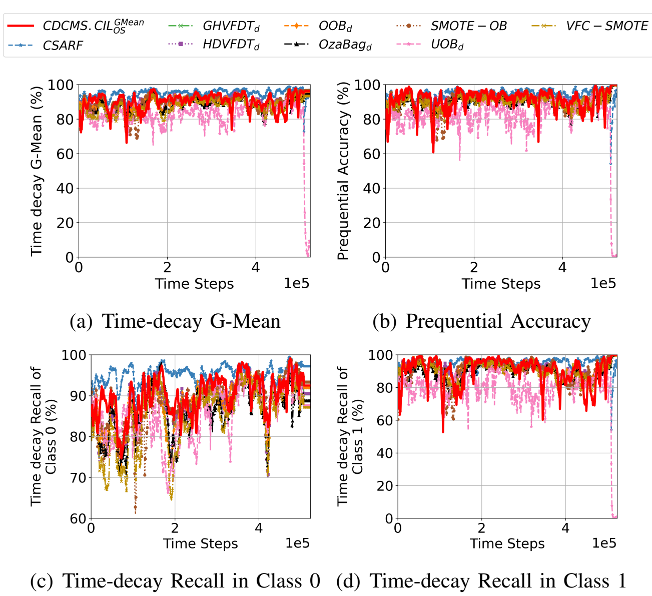
Fig. 40. Predictive Performance of Homo/Heterogeneous Approaches on Airlines



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

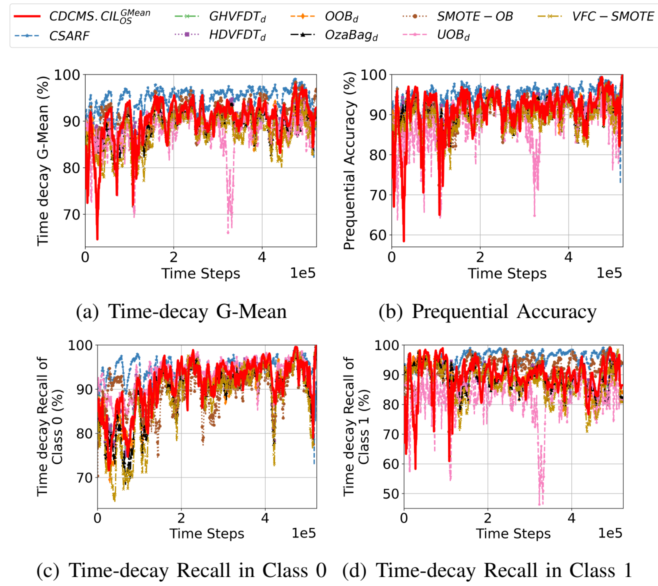
Fig. 41. Predictive Performance of Homo/Heterogeneous Approaches on NOAA



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

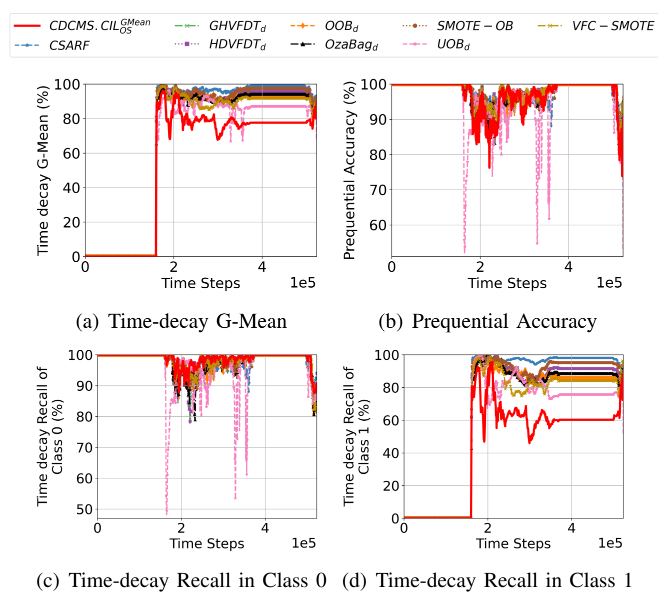
Fig. 42. Predictive Performance of Homo/Heterogeneous Approaches on Covtype$_{(c_1=1\text{-}6)}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

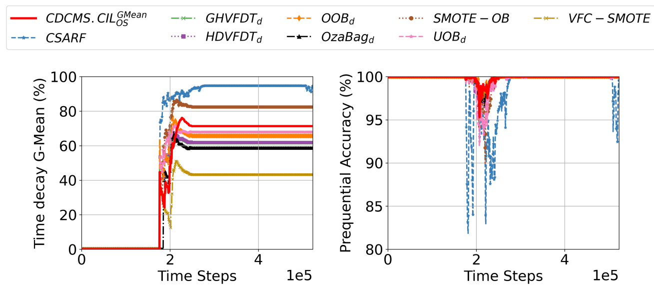
Fig. 43. Predictive Performance of Homo/Heterogeneous Approaches on Covtype$_{(c_1=1)}$



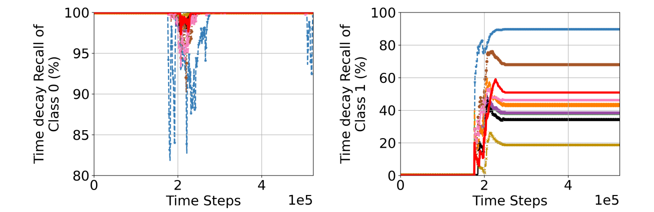
(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 44. Predictive Performance of Homo/Heterogeneous Approaches on Covtype$_{(c_1=2)}$

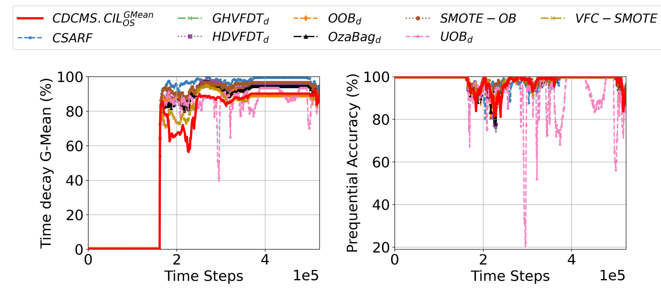


(a) Time-decay G-Mean (b) Prequential Accuracy

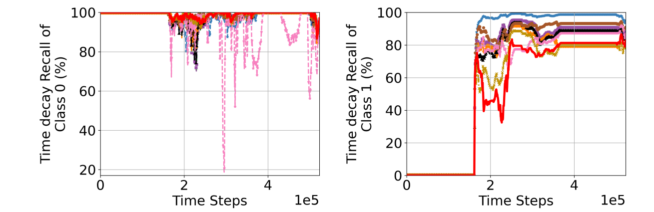


(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 45. Predictive Performance of Homo/Heterogeneous Approaches on Covtype$_{(c_1=3)}$

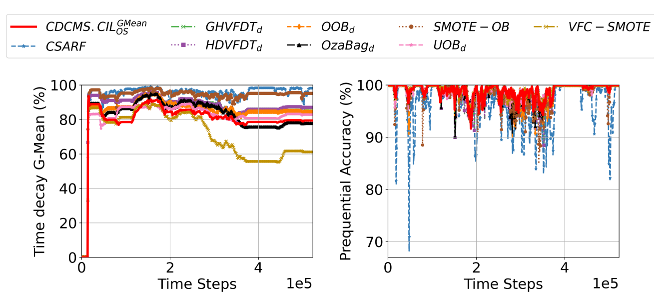


(a) Time-decay G-Mean (b) Prequential Accuracy

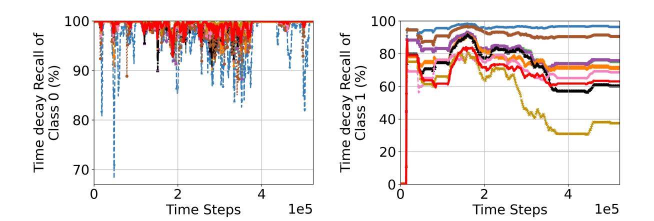


(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 47. Predictive Performance of Homo/Heterogeneous Approaches on Covtype$_{(c_1=5)}$

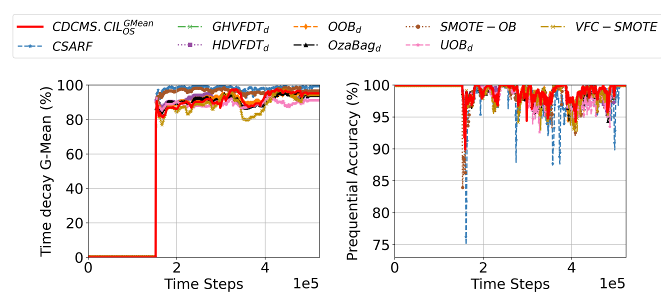


(a) Time-decay G-Mean (b) Prequential Accuracy

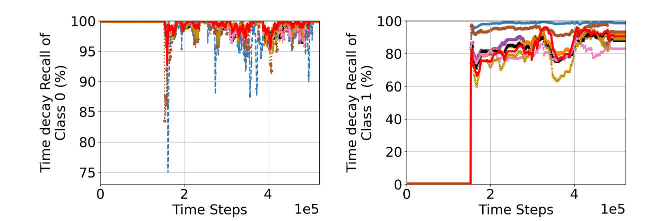


(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 46. Predictive Performance of Homo/Heterogeneous Approaches on Covtype$_{(c_1=4)}$

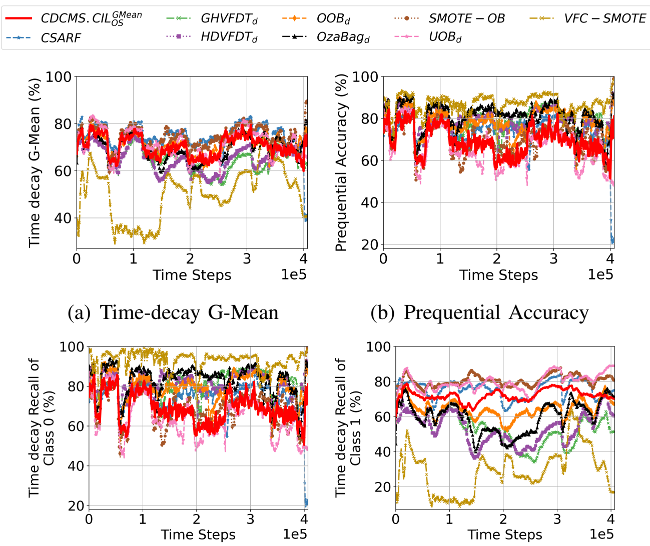


(a) Time-decay G-Mean (b) Prequential Accuracy



(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

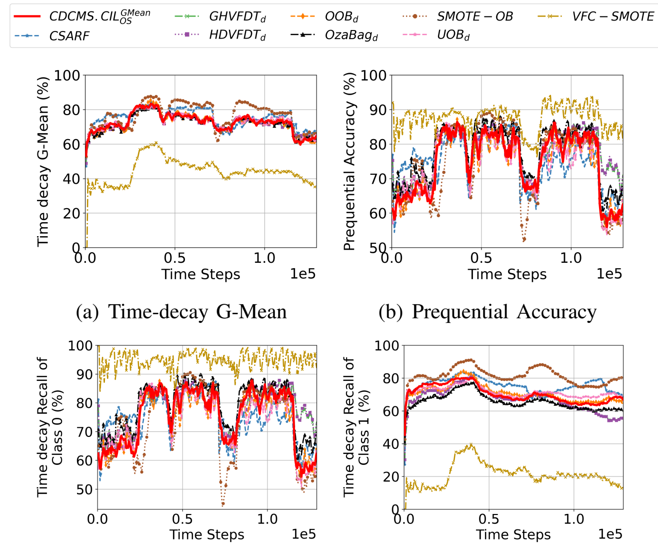
Fig. 48. Predictive Performance of Homo/Heterogeneous Approaches on Covtype$_{(c_1=6)}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

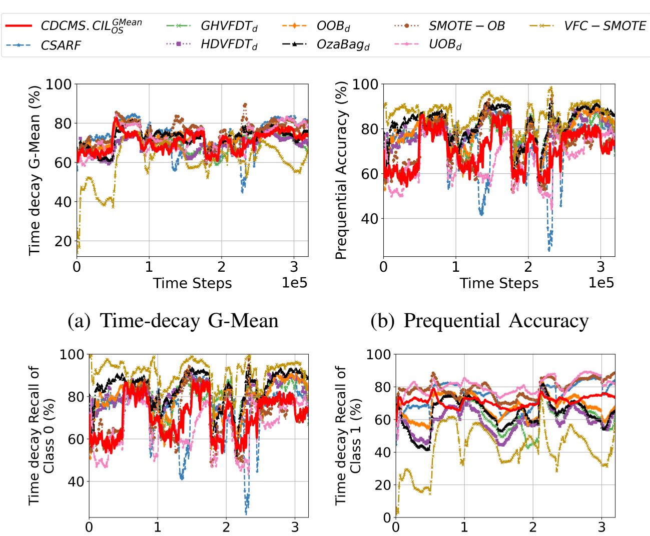
Fig. 49. Predictive Performance of Homo/Heterogeneous Approaches on INSECTS$^{\text{incremental}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

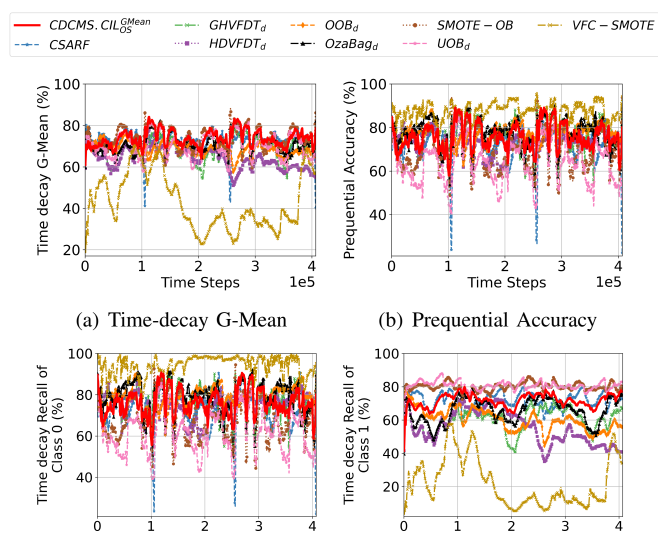
Fig. 51. Predictive Performance of Homo/Heterogeneous Approaches on INSECTS$^{\text{incremental}}_{\text{gradual}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

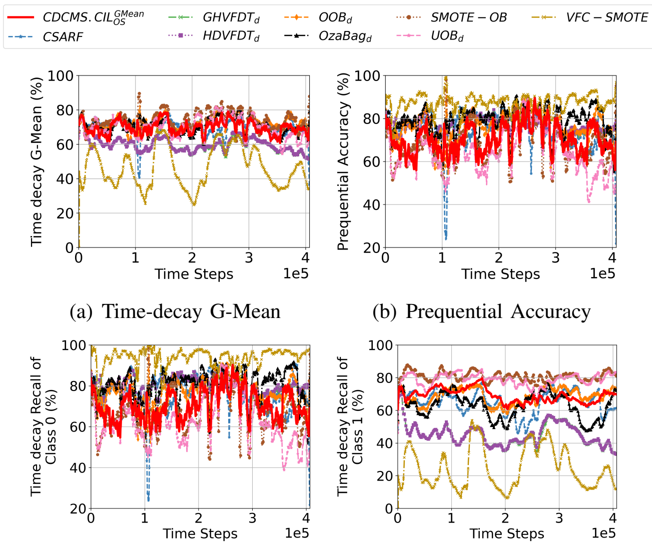
Fig. 50. Predictive Performance of Homo/Heterogeneous Approaches on INSECTS$_{\text{abrupt}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

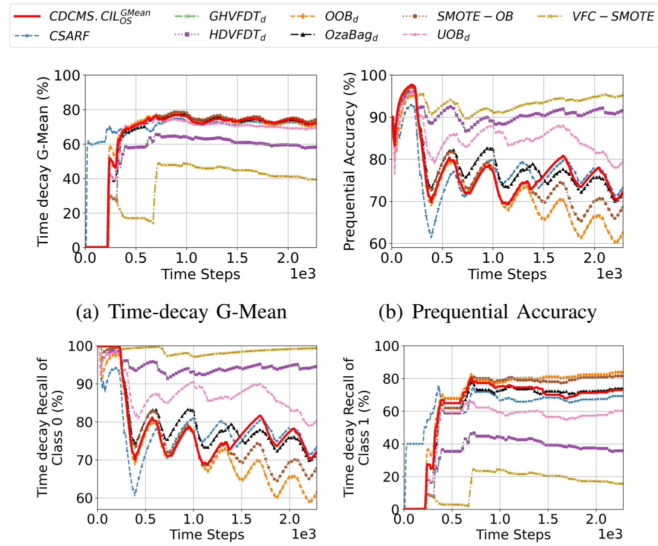
Fig. 52. Predictive Performance of Homo/Heterogeneous Approaches on INSECTS$^{\text{incremental}}_{\text{abrupt recurring}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

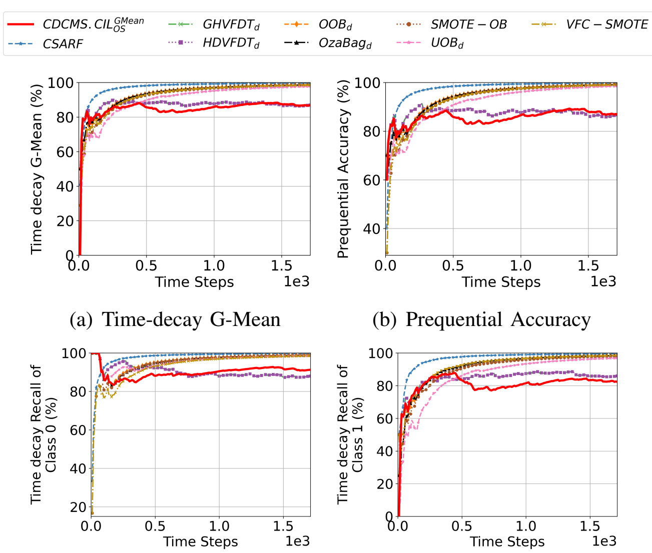
Fig. 53. Predictive Performance of Homo/Heterogeneous Approaches on INSECTS$_{\text{recurring}}^{\text{incremental}}$



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

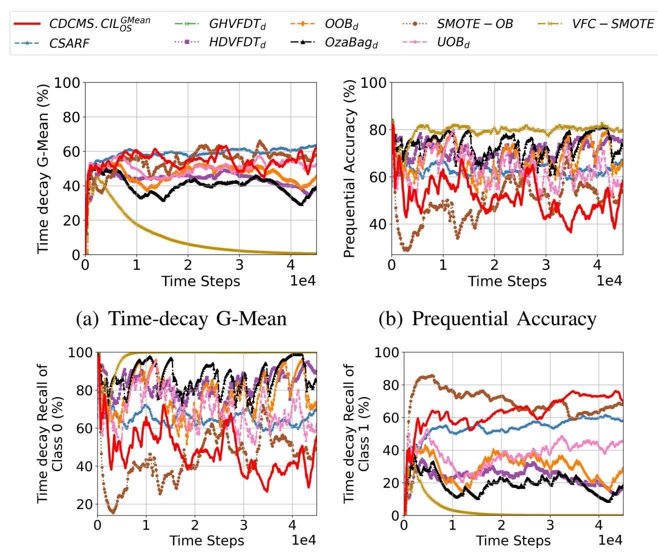
Fig. 55. Predictive Performance of Homo/Heterogeneous Approaches on Ozone



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

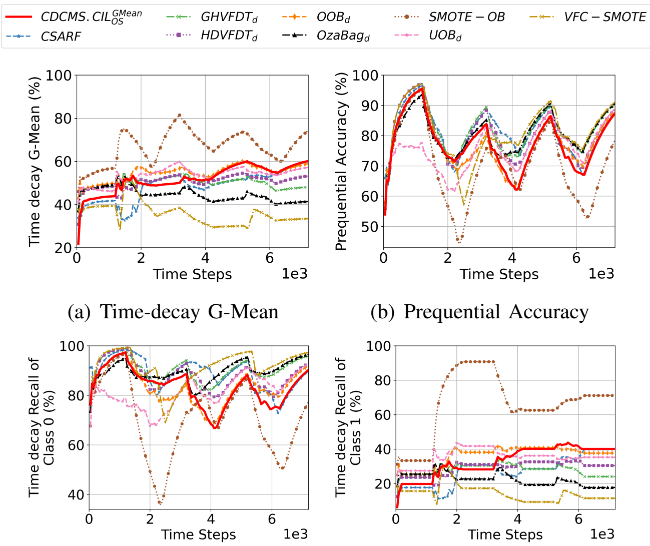
Fig. 54. Predictive Performance of Homo/Heterogeneous Approaches on Luxembourg



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

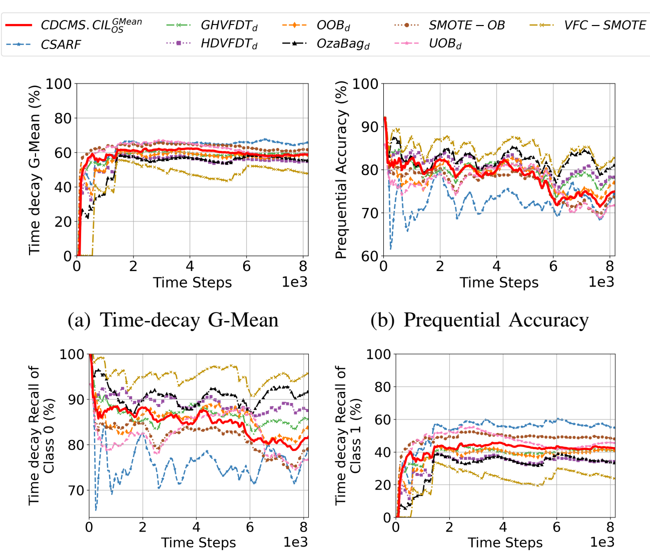
Fig. 56. Predictive Performance of Homo/Heterogeneous Approaches on PAKDD-2009



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 57. Predictive Performance of Homo/Heterogeneous Approaches on Amazon



(a) Time-decay G-Mean (b) Prequential Accuracy

(c) Time-decay Recall in Class 0 (d) Time-decay Recall in Class 1

Fig. 58. Predictive Performance of Homo/Heterogeneous Approaches on Twitter