SMOClust: Synthetic Minority Oversampling based on Stream Clustering for Evolving Data Streams - Supplementary Document

Chun Wai Chiu^{1^*} and Leandro L. Minku^{2^*}

 ^{1*}School of Computer Science and Mathematics, Keele
 University, Keele, Staffordshire, ST5 5BG, United Kingdom.
 ^{2*}School of Computer Science, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom.

*Corresponding author(s). E-mail(s): c.chiu@keele.ac.uk; l.l.minku@bham.ac.uk;

1 General

This is a supplementary document to the paper of "SMOClust: Synthetic Minority Oversampling based on Stream Clustering for Evolving Data Streams" and it is organised as follows:

- Section 2 presents the pseudo-code of Synthetic Minority Oversampling based on stream Clustering (SMOClust) written in a lower level of abstraction.
- Sections 3 and 4 presents the comprehensive results of the predictive performance of approaches on artificial data streams and real-world data streams respectively.

2 Proposed Approach

This section presents the pseudo-code of SMOClust written in a lower level of abstraction. Algorithm 1 presents the pseudo-code over-viewing SMOClust. Algorithm 2 presents the pseudo-code of this method. The details of generating a synthetic minority class example using micro-clusters can be described as

follows. Algorithm 3 presents the details of how to combine a set of microclusters into one. Algorithm 4 presents the algorithm of sampling from a hypersphere based on a skewed Gaussian where the maximum of the probability density function is predefined.

Algorithm 1 Synthetic Minority Oversampling based on Stream Clustering - SMOClust

```
Hyper-parameters: Base Learner(b), Stream Clustering Method(sc),
                                                                                     Class Size
                                                                                                   Fad-
ing Factor(\theta), Gaussian Noise Variance(v), Categorical Change Probability(P_c), k-Nearest
neighbour(k), Data Stream(S)
Variables: Base Learner(B), Stream Clustering Methods array(SC[])
1: SC[] \leftarrow createStreamClusteringMethods(sc, 2) \triangleright "2" refers to the number of SC to create.
2: for s_t \in S do
3:
       drift\_level \leftarrow DriftDetection(B, s_t)
       if drift\_level == DRIFT then
4:
5:
           B \leftarrow createNewBaseLearner(b); B.resetClassSize()
6:
       end if
7:
       B.trainOnInstance(s_t); B.updateClassSize(s_t, \theta)
8:
       last_inst[s_t.classValue()] \leftarrow s_t
                                                         \triangleright Store the last seen example of each class
       c_{maj} \leftarrow getMajorityClass(); c_{min} \leftarrow getMinorityClass()
Q٠
        while (B.rawClassSize(c_{min}) < B.rawClassSize(c_{maj})) and ((SC[0].isReady()) and
10:
   SC[1].isReady()) or last\_inst[c_{min}] \neq NULL) do
           if SC[0].isReady() and SC[1].isReady() then
11:
               mCluster_{anchor} \leftarrow weightedRandomDrawByAvgTimeStamp(SC[c_{min}])
12:
13.
               if mCluster_{anchor}.surroundedBySameClass(SC) then
                   synthInst^{Bin} \leftarrow genSynthInstBykNN(SC[c_{min}], mCluster_{anchor}, c_{min}, k)
14:
   \triangleright Alg. 2
15:
               else
                   synthInst^{Bin} \leftarrow genSynthInstByGauSampling(mCluster_{anchor})
16:
17:
               end if
               synthInst \leftarrow binaryToNominal(synthInst^{Bin}.copy())
18.
               SC[c_{min}].trainOnInstance(synthInst<sup>Bin</sup>.deleteClassAttribute())
19 \cdot
20:
               B.trainOnInstance(synthInst); B.updateClassSize(synthInst, \theta)
21:
           else
                                                                        \triangleright if last\_inst[c_{min}] \neq NULL
               synthInst \leftarrow addGaussianNoiseToInstance(last_inst[c_{min}], v, \dot{P_c})
22.
               synthInst^{Bin} \leftarrow nominalToBinary(synthInst.copy())
23.
               SC[c_{min}].trainOnInstance(synthInst<sup>Bin</sup>.deleteClassAttribute())
24:
25:
               B.trainOnInstance(synthInst); B.updateClassSize(synthInst, \theta)
26:
           end if
       end while

s^{Bin, noClass} \leftarrow nominalToBinary(s_t.copy())
27.
28:
        SC[s_t.classValue()].trainOnInstance(s_t^{Bin,noClass}.deleteClassAttribute())
29.
30: end for
```

Algorithm 2 Generate Synthetic Instance with k-NN Micro-Clusters

1: function GENSYNTHINSTBYKNN $(SC[c_{min}], mCluster_{anchor}, c_{min}, k)$

- 2: $kNNmClusters \leftarrow SC[c_{min}].getkNNmClusters(mCluster_{anchor})$
- 3: $sphere_cluster \leftarrow createSphereCluster(mCluster_{anchor}, kNNmClusters)$
- 4: $synthInst \leftarrow sphere_cluster.sample_around_target(anchor_mCluster.getCentre())$ Alg. 4
- 5: $synthInst.setClassValue(s_t.classValue())$
- 6: return synthInst
- 7: end function

Algorithm 3 Combining a set of micro-clusters into one
1: function COMBINE(mClusters[])
2: $dimensions \leftarrow mClusters[0].numOfDimensionss()$
3: for $i \in range(0mClusters.length)$ do
4: $all_centres[i] \leftarrow mClusters[i].getCentre()$
5: $all_weights[i] \leftarrow mClusters[i].getWeight()$
$6: \qquad all_radius[i] \leftarrow mClusters[i].getRadius()$
7: end for
8: $newCentre \leftarrow createArrayWithSize(dimensions)$
9: for $i \in range(0dimensions)$ do \triangleright Weighted sum of centres, by dimension.
10: $result_by_dim \leftarrow 0$
11: for $j \in range(0mClusters.length)$ do
$12: result_by_dim \leftarrow result_by_dim + all_centres[j][i] * all_weights[j]$
13: end for
$14: newCentre[i] = result_by_dim/sum(all_weights)$
15: end for
16: $r_n \leftarrow createArrayWithSize(all_radius.length)$
17: for $i \in range(0all_radius.length)$ do \triangleright Find the distance from <i>newCentre</i> to farthest
hull.
$18: \qquad distance_to_newCentre \leftarrow euclidean_distance(all_centres[i], newCentre)$
19: $r_n \leftarrow r_n \cup (all_radius[i] + abs(distance_to_newCentre))$
20: end for
21: $r_n \leftarrow r_n.sort(descending); new_radius \leftarrow r_n[0]$
22: return createMicroCluster(newCentre, newRadius)
23: end function

Algorithm 4 Sampling from a Hyper-Sphere by Skewed Gaussian with the Maximum of the Probability Density Function at a Designated Location

1:	function SAMPLE_AROUND_TARGET($\alpha^{(1)}$, sphere_cluster)	
2:	$\beta \leftarrow sphere_cluster.getCentre()$	
3:	$r \leftarrow sphere_cluster.getRadius()$	
4:	dimensions $\leftarrow \beta.numOfDimensions()$	
5:	$\delta \leftarrow createArrayWithSize(dimensions)$	
6:	$\gamma \leftarrow createArrayWithSize(dimensions)$	
7:	$\alpha^{(2)} \leftarrow sample_random_from_hypersphere(\alpha^{(1)}, 1)$	▷ By Muller's Method [?]
8:	$A \leftarrow 0; B \leftarrow 0; C \leftarrow 0$	
9:	for $i \in range(0dimensions)$ do	
10:	$\delta[i] \leftarrow \alpha^{(2)}[i] - \alpha^{(1)}[i]$	
11:	$\gamma[i] \leftarrow \beta[i] - \alpha^{(1)}[i]$	
12:	$A \leftarrow A + (\delta[i] * \delta[i])$	$\triangleright A = \sum_{i=0}^{n} \delta_i^2$
13:	$B \leftarrow B + (\delta[i] * \gamma[i])$	$\triangleright A = \sum_{i=0}^{n} \delta_i^2$ $\triangleright \sum_{i=0}^{n} \delta_i \gamma_i$
14:	$C \leftarrow C + (\gamma[i] * \gamma[i])$	$\triangleright \sum_{i=0}^{n} \gamma_i^2$
15:	end for	
16:	$B \leftarrow B * -2$	$\triangleright B = -2(\sum_{i=0}^{n} \delta_i \gamma_i)$
17:	$C \leftarrow C - (r * r)$	$\triangleright C = \left(\sum_{i=0}^{n} \gamma_i^2\right) - r^2$
18:	$\mathbf{return} \ (-B + sqrt(B * B - 4 * A * C))/(2 * A)$	$\triangleright \frac{-B + \sqrt{B^2 - 4AC}}{2A}$
19:	end function	

3 Results with Artificial Data Streams

This section resents the comprehensive results of the predictive performance of approaches on artificial data streams.

- Correspond to Figure 4 in the paper:
 - Figure 1 presents the difference in average G-Mean (based on thirty runs) of the compared approaches against SMOClust on **five**-dimensional class imbalanced artificial data streams.
 - Table 3 presents average G-Mean (based on thirty runs) of all approaches on **five** dimensional class imbalanced artificial data streams and the A12 effect size results of comparing existing approaches against SMOClust.
- Correspond to Figure 5 in the paper:
 - Figure 2 presents the difference in average G-Mean (based on thirty runs) of the compared approaches against SMOClust on **five**-dimensional severely class imbalanced artificial data streams.
 - Table 4 presents average G-Mean (based on thirty runs) of all approaches on **five**-dimensional severely class imbalanced artificial data streams and the A12 effect size results of comparing existing approaches against SMOClust.
- Correspond to Figure 14 in the paper:
 - Figure 3 presents the difference in average G-Mean (based on thirty runs) of the compared approaches against SMOClust on two-dimensional class imbalanced artificial data streams.
 - Table 5 presents average G-Mean (based on thirty runs) of all approaches on two-dimensional class imbalanced artificial data streams and the A12 effect size results of comparing existing approaches against SMOClust.
- Correspond to Figure 15 in the paper:
 - Figure 4 presents the difference in average G-Mean (based on thirty runs) of the compared approaches against SMOClust on **two**-dimensional severely class imbalanced artificial data streams.
 - Table 6 presents average G-Mean (based on thirty runs) of all approaches on two-dimensional severely class imbalanced artificial data streams and the A12 effect size results of comparing existing approaches against SMOClust.
- Correspond to Tables 8 in the paper:
 - Table 1 presents average G-Mean (based on thirty runs) of the approaches on the two-dimensional version of StaticIm1_Move7 stream and the A12 effect size results of comparing existing approaches against SMOClust.
- Correspond to Tables 9 in the paper:

- Table 2 presents average G-Mean (based on thirty runs) of the approaches on the two-dimensional version of StaticIm10_Rare100 stream and the A12 effect size results of comparing existing approaches against SMO-Clust.
- Correspond to Figure 6 in the paper:
 - Figure 5 compares the predictive performance of SMOClust with that of OOB_d , UOB_d , oOS_d , and $oUnderOverB_d$ in the StaticIm1_Move7 stream. The comparison is made over time steps in the median run¹ of the approaches.
- Correspond to Figure 10 in the paper:
 - Figure 6 compares the predictive performance of SMOClust with that of OOB_d , UOB_d , oOS_d , and $oUnderOverB_d$ in the StaticIm10_Rare100 stream. The comparison is made over time steps in the median run of the approaches.

 $^{^1\}mathrm{Median}$ run refers to the run that leads to the median of predictive performances averaged across time steps.

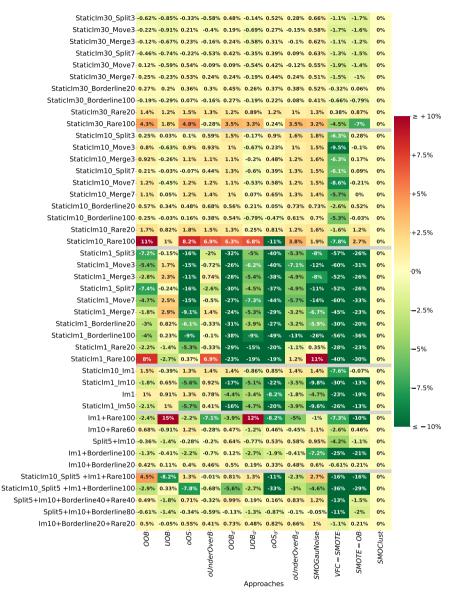


Fig. 1: Difference in Average G-Mean Against SMOClust on Class Imbalanced Artificial Data Streams Based on 30 Runs (Green cells indicate SMOClust performed better; Red cells indicate SMOClust performed worse; Grey horizontal lines separate different groups of data streams, i.e., $StaticIm\{30/10/1\}$, Imbalance Ratio Drift, Double Factor, and Complex Factor)

Streams

Chabialas E. Caliba	0.050/			0.010/		0.070/					0.000/		
StaticIm5_Split3 StaticIm5 Move3	0.05%	0.35%	-0.72% 0.51%	0.91%	1% 1.2%	0.07% -0.75%	-1.1%	1.9% 1.6%	1.9% 1.9%	-16% -23%	-0.93% -0.95%	0% 0%	
StaticIm5_Moves		-0.06%	1.2%	2%	1.5%	-0.11%	-5.1%	2%	2.2%	-23%	-0.2%	0%	
StaticIm5_Merges		0.24%	-0.94%	0.83%	0.47%	-0.38%	-2.3%	2%	1.9%	-17%	-0.2%	0%	
StaticIm5_Split7		-0.94%	0.52%	1.6%	0.81%	-1.3%	-3.7%	1.4%	1.6%	-22%	-1.5%	0%	
StaticIm5 Merge7		0.19%	1.2%	2%	1.2%	0.01%	-2.8%	1.8%	1.9%	-15%	-0.84%	0%	
StaticIm5 Borderline20		0.23%	0.07%	0.73%	-0.3%	0.04%	-2.1%	0.57%	0.56%	-6.6%	0.18%	0%	
StaticIm5 Borderline100		-0.46%	-0.33%	0.54%	-0.27%	-1.1%	-8.8%	0.7%	0.57%	-17%	-0.32%	0%	
StaticIm5_Borderine100		0.14%	1.4%	1.3%	0.58%	-0.6%	-1.2%	0.89%	1.5%	-5.4%	0.67%	0%	
StaticIm5 Rare100		-0.89%	7.6%	9.4%	6.3%	6.1%	-13%	0.88%	-0.3%	-16%	0.63%	0%	
StaticIm3 Split3		0.77%	-2.7%	0.61%	-2.3%	0.33%	-7.1%	1.7%	1.4%	-27%	-4.3%	0%	
StaticIm3 Move3		-0.23%	-1.4%	1.9%	-0.24%	-0.92%	-12%	1.6%	1.6%	-35%	-5.1%	0%	
StaticIm3 Merge3		0.33%	0.16%	2.5%	0.08%	-0.15%	-9.4%	2.2%	2.4%	-26%	-2.5%	0%	$\geq +10^{\circ}$
StaticIm3 Split7		0.9%	-3.3%	0.61%	-3.5%	-0.21%	-9.9%	1.6%	1.6%	-26%	-4.4%	0%	
StaticIm3 Move7		-0.64%	-1.2%	1.8%	-0.77%	-1.3%	-11%	1.2%	0.9%	-34%	-7%	0%	
StaticIm3_Move7		0.18%	0.02%	2%	-0.44%	-0.18%	-8.9%	1.2%	1.1%	-23%	-4.2%	0%	
StaticIm3 Borderline20		0.18%	-0.69%	0.57%	-3.8%	-0.23%	-5.2%	0.17%	-0.13%	-23%	-4.2%	0%	. 7 . 5 0/
													+7.5%
StaticIm3_Borderline100		-0.61%	-1.3%	0.57%	-3.8%	-1.3%	-24%	0.61%	-0.51%	-30%	-2.8%	0%	
StaticIm3_Rare20		-0.25%	0.63%	1.2%	-2.2%	-1.5%	-2.3%	0.99%	2.1%	-8.2%	-0.61%	0%	
StaticIm3_Rare100		2.3%	6.2%	9.7%	5%	2.3%	-6.9%	2.7%	1.9%	-24%	-16%	0%	
StaticIm1_Split3		0.07%	-15%	-2.2%	-29%	-5.3%	-37%	-5.1%	-8.7%	-54%	-26%	0%	+5%
StaticIm1_Move3		1.6%	-18%	-1.6%	-19%	-4.4%	-43%	-8.4%	-13%	-63%	-34%	0%	
StaticIm1_Merge3		2%	-13%	0.68%	-23%	-4.4%	-38%	-5.3%	-7.7%	-53%	-28%	0%	
StaticIm1_Split7		-0.87%	-17%	-3%	-31%	-4.9%	-38%	-5.7%	-11%	-52%	-25%	0%	
StaticIm1_Move7		2.3%	-16%	-0.88%	-20%	-7.7%	-42%	-8.1%	-12%	-61%	-34%	0%	
StaticIm1_Merge7	-3%	2.3%	-9.5%	1%	-16%	-5.3%	-30%	-4.7%	-6.6%	-46%	-24%	0%	+2.5%
StaticIm1_Borderline20	-3.5%	0.59%	-6.3%	-0.58%	-24%	-4.5%	-25%	-3.7%	-5.7%	-31%	-19%	0%	
StaticIm1_Borderline100	-4.3%	0.18%	-8.9%	-0.27%	-38%	-7.2%	-50%	-16%	-26%	-55%	-33%	0%	
ച StaticIm1_Rare20		-1.5%	-5.6%	-0.65%	-22%	-13%	-23%	-1.9%	-0.58%	-29%	-22%	0%	
StaticIm1_Rare20 StaticIm1_Rare100 StaticIm07_Split3 StaticIm07_Move3	7.3%	-2.6%	0.04%	6.1%	-18%	-16%	-19%	0.58%	9.8%	-40%	-30%	0%	.0%
StaticIm07_Split3	-12%	0.41%	-24%	-6.1%	-44%	-9.2%	-53%	-14%	-19%	-67%	-45%	0%	.0%
សី StaticIm07 Move3	-13%	3.3%	-30%	-6.1%	-43%	-10%	-53%	-17%	-23%	-70%	-51%	0%	
StaticIm07 Merge3	-8%	4.3%	-21%	-1.7%	-40%	-10%	-50%	-11%	-17%	-64%	-45%	0%	
StaticIm07 Split7	-13%	-0.05%	-25%	-6.6%	-46%	-11%	-51%	-13%	-17%	-63%	-43%	0%	
StaticIm07 Move7	-13%	4%	-28%	-5.4%	-42%	-11%	-51%	-18%	-25%	-69%	-51%	0%	-2.5%
StaticIm07 Merge7	-7.3%	3.9%	-18%	-1.5%	-38%	-7%	-41%	-10%	-13%	-54%	-40%	0%	
StaticIm07 Borderline20	-5.8%	1.3%	-11%	-1.8%	-39%	-6.5%	-47%	-5.8%	-15%	-42%	-36%	0%	
StaticIm07 Borderline100		0.83%	-15%	-1.8%	-53%	-16%	-62%	-26%	-36%	-65%	-50%	0%	
StaticIm07 Rare20		-1.5%	-10%	-2%	-41%	-21%	-47%	-4.8%	-5.4%	-40%	-38%	0%	
StaticIm07 Rare100	6.6%	-3.6%	-3.3%	4.7%	-36%	-22%	-33%	0.68%	8.1%	-46%	-38%	0%	-5%
StaticIm05 Split3		2.1%	-35%	-12%	-58%	-17%	-63%	-26%	-29%	-75%	-63%	0%	
StaticIm05 Move3		7%	-40%	-12%	-57%	-23%	-60%	-31%	-33%	-75%	-64%	0%	
StaticIm05 Merge3		6.3%	-32%	-7.2%	-58%	-20%	-59%	-24%	-31%	-73%	-61%	0%	
StaticIm05 Split7		1.5%	-34%	-12%	-56%	-21%	-60%	-22%	-27%	-69%	-60%	0%	-7.5%
StaticIm05 Move7	-22%	6.9%	-39%	-11%	-51%	-27%	-59%	-30%	-35%	-73%	-65%	0%	1.570
StaticIm05 Merge7	-14%	5.7%	-27%	-6%	-48%	-17%	-56%	-16%	-22%	-64%	-56%	0%	
StaticIm05_Borderline20		2.4%	-17%	-3.6%	-63%	-15%	-57%	-17%	-24%	-53%	-54%	0%	
StaticIm05 Borderline100		2.1%	-23%	-4%	-65%	-33%	-66%	-38%	-43%	-76%	-68%	0%	
StaticIm05 Rare20		-0.66%	-17%	-3.8%	-58%	-31%	-53%	-9.7%	-13%	-52%	-56%	0%	$\leq -10^{\circ}$
StaticIm05 Rare100		-4.5%	-8.7%	1.8%	-50%	-31%	-37%	-7.2%	3.9%	-54%	-48%	0%	
StaticIm03 Split3		7%	-47%	-23%	-67%	-36%	-67%	-39%	-41%	-80%	-74%	0%	
StaticIm03_Splits		14%	-47%	-25%	-57%	-36%	-65%	-39% -40%	-41% -40%	-80%	-74% -70%	0%	
StaticIm03_Moves		14%	-33% -45%	-25% -17%	-57% -58%	-35% -29%	-62%	-40%	-40% -41%	-75%	-70% -70%	0%	
StaticInf05_Merges StaticIm03 Split7		6%	-45% -48%	-17%	-58% -65%	-29% -32%	-62%	-34% -39%	-41% -41%	-75% -78%	-70% -72%	0%	
StaticIm03_Split7		15%	-48% -52%	-23% -22%	-65% -57%	-32% -47%	-65%	-39% -42%	-41% -41%	-78% -75%	-72%	0%	
StaticIm03_Move7 StaticIm03_Merge7	-35% -22%	15%	-52% -39%	-22% -12%	-57% -55%	-42% -26%	-65% -61%	-42% -25%	-41% -28%	-75% -73%	-70% -68%	0%	
StaticIm03_Borderline20		4.3%	-37% -39%	-10% -12%	-69% -70%	-33% -45%	-72% -72%	-35% -51%	-49% -50%	-76% -86%	-77% -82%	0%	
StaticIm03_Borderline100		4.6%										0%	
StaticIm03_Rare20		0.94%	-36%	-10%	-70%	-45%	-70%	-30%	-36%	-73%	-77%	0%	
StaticIm03_Rare100		-6%	-19%	-4.1%	-54%	-40%	-48%	-17%	-7.5%	-59%	-57%	0%	
	OOB.	UOB.	00S-	oUnderOverB	Bd	Bd	005d	B	se	- SMOTE	SMOTE – OB	SMOClust-	
	ŏ	ษ	ŏ	Ne	OOBd	UOBd	8	oUnderOverB _d	SMOGauNoise	Q	ĭ	C	
				2	2	~	-	Ó	Ine	SV	E	Q	
				de				Jer	ğ	<u> </u>	10	SV	
				Ľ,				Щ	00	ñ	SN		
				0				or	SI	2			
						Appro	aches						

Fig. 2: Difference in Average G-Mean Against SMOClust on Class Imbalanced Artificial Data Streams Based on 30 Runs (Green cells indicate SMOClust performed better; Red cells indicate SMOClust performed worse; Grey horizontal lines separate different groups of data streams, i.e., StaticIm $\{5/3/1/07/05/03\}$)

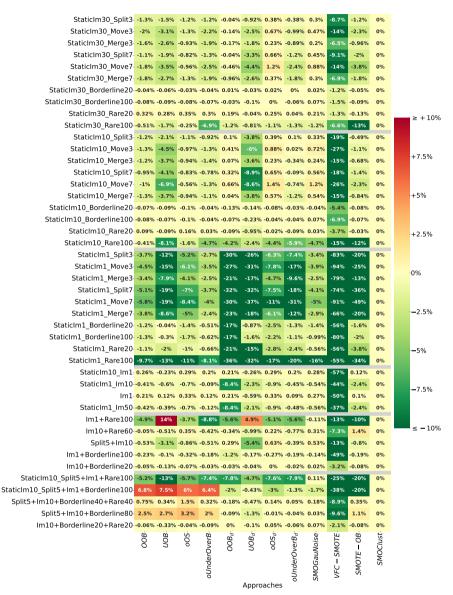


Fig. 3: Difference in Average G-Mean Against SMOClust on Two-Dimensional Class Imbalanced Artificial Data Streams Based on 30 Runs (Green cells indicate SMOClust performed better; Red cells indicate SMOClust performed worse; Grey horizontal lines separate different groups of data streams, i.e., StaticIm{30/10/1}, Imbalance Ratio Drift, Double Factor, and Complex Factor)

StaticIm5_Split3 StaticIm5 Move3		-3.3% -6.7%	-1.3% -0.92%	-0.77% -1%	-0.6%	-10% -8.6%	0.03%	0.28%	0.21%	-34% -45%	-0.45%	0% 0%		
StaticIm5_Moves		-6.7%	-0.92%	-1%	0.02% -0.18%	-8.6%	0.12%	-0.31%	0.7%	-45% -28%	-0.82% -0.65%	0%		
StaticIm5_Merges		-6.7%	-0.93%	-1.2%	-0.65%	-14%	0.12%	-0.13%	0.26%	-28%	-1.4%	0%		
StaticIm5_Spirt7		-9.2%	-0.79%	-1%	0.02%	-16%	1%	-0.6%	1.1%	-47%	-2.6%	0%		
StaticIm5_Move7		-4.3%	-0.96%	-0.98%	-0.26%	-4.3%	0.43%	-1%	0.45%	-27%	-1.2%	0%		
StaticIm5 Borderline20			-0.23%	-0.07%	-1.4%	-0.23%	-0.27%	-0.1%	-0.1%	-9.9%	-0.11%	0%		
StaticIm5 Borderline100		-0.1%	-0.19%	-0.06%	-1.1%	-0.46%	-0.18%	-0.1%	0.06%	-15%	-0.1%	0%		
StaticIm5 Rare20			0.09%	-0.14%	-1.1%	-2.2%	-0.21%	-0.34%	0.12%	-7.7%	-0.24%	0%		
StaticIm5 Rare100		-16%	-3.3%	-3.7%	-8.3%	-4%	-8.6%	-12%	-9.4%	-27%	-10%	0%		
StaticIm3 Split3		-4.3%	-1.7%	-0.95%	-2.5%	-13%	-0.57%	-0.62%	-0.17%	-50%	-0.85%	0%		
StaticIm3 Move3		-7.9%	-1.5%	-1.1%	-1.5%	-9.6%	-0.27%	-0.9%	0.17%	-67%	-1.5%	0%		
StaticIm3 Merge3		-5.6%	-1.3%	-1.2%	-1.1%	-5.8%	-0.4%	-0.93%	-0.06%	-44%	-0.97%	0%	≥	+ 109
StaticIm3 Split7		-11%	-2%	-1.1%	-4.4%	-16%	-0.89%	-1.1%	-0.14%	-45%	-3.4%	0%		
StaticIm3 Move7		-12%	-1.8%	-1.2%	-3.5%	-17%	-0.48%	-2%	0.24%	-65%	-4%	0%		
StaticIm3 Merge7		-5.3%	-1.3%	-1%	-2.2%	-5.6%	-0.35%	-1.5%	0.09%	-38%	-1.9%	0%		
StaticIm3 Borderline20		-0.15%	-0.46%	-0.15%	-4.5%	-0.29%	-0.71%	-0.25%	-0.29%	-16%	-0.23%	0%	+7	7.5%
StaticIm3 Borderline100		-0.23%	-0.45%	-0.14%	-3.5%	-0.59%	-0.45%	-0.22%	-0.1%	-33%	-0.18%	0%		1070
StaticIm3 Rare20	-0.24%	-0.84%	-0.15%	-0.26%	-3.8%	-3.9%	-0.6%	-0.6%	0.09%	-17%	-0.54%	0%		
StaticIm3 Rare100		-18%	-6.1%	-4.4%	-13%	-12%	-12%	-17%	-13%	-39%	-25%	0%		
StaticIm1 Split3		-11%	-5.3%	-2.8%	-20%	-25%	-6.7%	-8.5%	-3%	-84%	-20%	0%		
StaticIm1 Move3	-4.3%	-15%	-5.8%	-3.2%	-27%	-31%	-7.8%	-13%	-3.7%	-94%	-30%	0%	+5	5%
StaticIm1 Merge3		-8.5%	-3.8%	-2.4%	-18%	-19%	-4.8%	-7.6%	-2.4%	-81%	-14%	0%		
StaticIm1 Split7	-5.1%	-19%	-7.2%	-3.7%	-33%	-29%	-7.5%	-16%	-3.6%	-74%	-35%	0%		
StaticIm1 Move7		-19%	-8.2%	-3.8%	-32%	-36%	-10%	-23%	-4.9%	-92%	-49%	0%		
StaticIm1 Merge7	-3.5%	-8.3%	-4.6%	-2.1%	-19%	-16%	-6%	-11%	-2.8%	-67%	-19%	0%	+7	2.5%
StaticIm1_Borderline20	-1.2%	-0.06%	-1.6%	-0.61%	-24%	-1.1%	-2.7%	-1.5%	-1.5%	-55%	-2%	0%		
StaticIm1 Borderline100		-0.35%	-1.7%	-0.67%	-20%	-1.7%	-2.5%	-1.1%	-1.1%	-80%	-2.4%	0%		
 StaticIm1 Rare20 	-1.1%	-1.9%	-1.1%	-0.77%	-21%	-15%	-2.6%	-2.5%	-0.25%	-57%	-4.5%	0%		
StaticIm1_Rare20 StaticIm1_Rare100	-9.6%	-15%	-11%	-8.2%	-41%	-33%	-17%	-19%	-16%	-55%	-34%	0%		,
	-5.7%	-13%	-7.5%	-4.2%	-34%	-39%	-8.9%	-14%	-4.7%	-90%	-38%	0%	.0%	6
ភ័ StaticIm07_Move3	-6.2%	-16%	-8.6%	-4.6%	-44%	-52%	-12%	-25%	-6.7%	-94%	-57%	0%		
StaticIm07_Merge3		-9.4%	-5.3%	-3.2%	-31%	-24%	-7.8%	-16%	-4.2%	-89%	-28%	0%		
StaticIm07_Split7		-19%	-9%	-4.6%	-44%	-38%	-11%	-25%	-4.7%	-78%	-47%	0%		
StaticIm07_Move7		-18%	-12%	-5%	-46%	-47%	-14%	-40%	-7%	-91%	-68%	0%	-2	2.5%
StaticIm07_Merge7		-8.8%	-6.8%	-2.9%	-32%	-21%	-8.7%	-21%	-4.4%	-74%	-31%	0%		
StaticIm07_Borderline20		0.11%	-2.2%	-0.84%	-22%	-2.1%	-4.3%	-2.6%	-2.4%	-71%	-5.1%	0%		
StaticIm07_Borderline100		-0.3%	-2.6%	-1%	-23%	-3.4%	-4.9%	-2.6%	-2.1%	-87%	-7.6%	0%		
StaticIm07_Rare20		-2.2%	-1.7%	-0.95%	-30%	-18%	-4.4%	-3%	-0.89%	-72%	-12%	0%		5%
StaticIm07_Rare100		-12%	-11%	-8%	-41%	-32%	-20%	-21%	-14%	-60%	-38%	0%	-	,,0
StaticIm05_Split3	-7%	-16%	-9.1%	-4.8%	-42%	-45%	-13%	-20%	-5.9%	-91%	-54%	0%		
StaticIm05_Move3		-14%	-11%	-5.7%	-50%	-61%	-16%	-28%	-7.4%	-92%	-71%	0%		
StaticIm05_Merge3		-8.5%	-6.9%	-4.1%	-37%	-31%	-11%	-18%	-5.1%	-92%	-41%	0%		
StaticIm05_Split7	-8%	-17%	-10%	-4.8%	-48%	-44%	-13%	-24%	-4.1%	-80%	-55%	0%	-7	7.5%
StaticIm05_Move7		-13%	-14%	-4.9%	-49%	-65%	-19%	-46%	-6.2%	-85%	-75%	0%		
StaticIm05_Merge7		-7.1%	-8.7%	-3.5%	-42%	-28%	-12%	-25%	-5.5%	-80%	-41%	0%		
StaticIm05_Borderline20		0.56%	-2.6%	-0.9%	-39%	-3.7%	-6.6%	-4.3%	-3.1%	-84%	-12%	0%		
StaticIm05_Borderline100		-0.01%	-3.8%	-1.4%	-41%	-6.3%	-8.4%	-5.4%	-3.4%	-91%	-21%	0%	≤	- 10%
StaticIm05_Rare20		-2.3%	-2.2%	-1.1%	-48%	-21%	-6.3%	-4.2%	-0.91%	-83%	-21%	0% 0%	-	_ , ,
StaticIm05_Rare100		-10%	-12%	-7.6%	-51%	-34%	-22%	-23%	-13%	-65%	-43%			
StaticIm03_Split3		-14%	-11%	-4.8%	-54%	-47%	-20%	-23%	-5.4%	-87%	-66%	0%		
StaticIm03_Move3		-6.7%	-13%	-4.9% -4.8%	-57% -46%	-70% -38%	-25% -20%	-38% -25%	-6.4%	-84%	-77%	0%		
StaticIm03_Merge3		-6.3%	-9.7%						-6.8%	-91%	-61%	0%		
StaticIm03_Split7 StaticIm03 Move7		-10% -1.7%	-12% -16%	-4.7% -4.4%	-56% -49%	-44% -64%	-21% -29%	-30% -42%	-3.5% -4.4%	-78% -74%	-63% -70%	0% 0%		
StaticIm03_Move7 StaticIm03_Merge7														
StaticIm03_Merge7 StaticIm03 Borderline20		-2.2% 1.9%	-11% -3.7%	-4% -0.98%	-47% -41%	-32% -8.9%	-22% -11%	-25% -8.1%	-5.3% -5.2%	-82% -93%	-58% -35%	0% 0%		
StaticIm03_Borderline100		1.9%	-3.7% -6.6%	-0.98%	-41% -49%	-8.9% -20%	-11%	-8.1%	-5.2% -9.9%	-93% -95%	-35% -49%	0%		
StaticIm03_Bordenme100 StaticIm03_Rare20		-1.7%	-0.6%	-2%	-49% -50%	-20% -29%	-22% -13%	-4.4%	-9.9%	-95% -90%	-49% -46%	0%		
StaticIm03_Rare20 StaticIm03_Rare100		-1.7%	-3.1%	-1.1%	-50% -51%	-29% -37%	-13% -27%	-4.4%	-1.8%	-90% -67%	-46% -53%	0% 0%		
Staticimu3_Kare100														
	OOB.	UOB.	00S	oUnderOverB	00B _d	UOBd	005 _d	'n,	ise	- SMOTE	- OB	SMOClust-		
	0	D	0	õ	б	ы	8	ve	No.	MC	1	DC		
				Jua				Q	au	ŝ	SMOTE	MC		
				pu,				эрц	90	<u>,</u>	MC	S		
				Do				oUnderOverB _d	SMOGauNoise	ΥF	S			
						Appro	aches	0	01					

Fig. 4: Difference in Average G-Mean Against SMOClust on Two-Dimensional Class Imbalanced Artificial Data Streams Based on 30 Runs (Green cells indicate SMOClust performed better; Red cells indicate SMOClust performed worse; Grey horizontal lines separate different groups of data streams, i.e., StaticIm $\{5/3/1/07/05/03\}$

Table 1: 30 Runs Average G-Mean on Two-Dimensional Version of Representative Artificial Data Streams where SMOClust Performed Better (A12 SMOClust vs Others)

Stream	OOB	UOB	oOS	oUnder- OverB	OOB_{d}	UOB _d
$StaticIm1_Move7$	82.11%[-b]	$76.3\%[ext{-b}]$	79.46%[-b]	85.26%[-b]	53.45%[-b]	56.94%[-b]
Stream	$\mathrm{oOS}_{\mathrm{d}}$	oUnder- OverB _d	SMO- GauNoise	VFC- SMOTE	SMOTE- OB	SMOClust
$StaticIm1_Move7$	76.88%[-b]	45.12%[-b]	82.94%[-b]	1.04%[-b]	33.09%[-b]	91.23%

⁻ Based on the average G-Mean, cells are highlighted in lime / light grey when SMOClust performed better than the corresponding approach and cells are highlighted in orange / dark grey cells when SMOClust performed worse than the corresponding approach. The colour intensity scales with the absolute difference in average G-Mean between the SMOClust and the approach of the column and the intensity reaches the maximum when such difference is $\geq 10\%$.

⁻ Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against SMOClust. Presence/absence of the sign "-" in the effect size means that the corresponding approach was worse/better than SMOClust.

Table 2: 30 Runs Average G-Mean on Two-Dimensional Version of Representative Artificial Data Streams where SMOClust Performed Worse (A12 SMOClust vs Others)

Stream	OOB	UOB	oOS	oUnder- OverB	OOB_{d}	UOB _d
$StaticIm10_Rare100$	70.61%[b]	63.65%[-b]	69.14%[-b]	68.19%[-b]	65.49%[-b]	68.17%[-b]
Stream	$\mathrm{oOS}_{\mathrm{d}}$	oUnder- OverB _d	SMO- GauNoise	VFC- SMOTE	SMOTE- OB	SMOClust
StaticIm10_Rare100	65.04%[-b]	64.98%[-b]	64.56%[-b]	54.64%[- b]	58.98%[-b]	70.32%

⁻ Based on the average G-Mean, cells are highlighted in lime / light grey when SMOClust performed better than the corresponding approach and cells are highlighted in orange / dark grey cells when SMOClust performed worse than the corresponding approach. The colour intensity scales with the absolute difference in average G-Mean between the SMOClust and the approach of the column and the intensity reaches the maximum when such difference is $\geq 10\%$.

Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against SMOClust. Presence/absence of the sign "-" in the effect size means that the corresponding approach was worse/better than SMOClust.

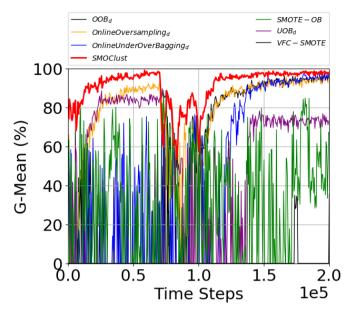


Fig. 5: Periodic Class Balanced Holdout Test G-Mean Against Time Steps in Two-Dimensional StaticIm1_Move7

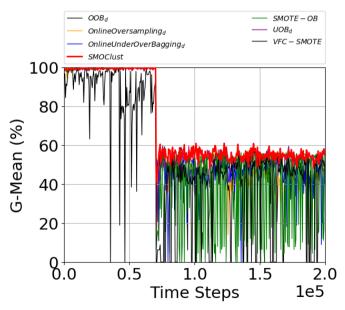


Fig. 6: Periodic Class Balanced Holdout Test G-Mean Against Time Steps in Two-Dimensional StaticIm10_Rare100

		0						· · · · · · · · · · · · · · · · · · ·			/	
Stream	OOB	UOB	oOS	oUnder- OverB	OOB_{d}	UOB_d	$\mathrm{oOS}_{\mathrm{d}}$	oUnder- OverB _d	SMO- Gau- Noise	VFC- SMOTE	SMOTE- OB	SMO- Clust
					StaticI	m30			110100			
Split3	96.56%[-b]	96.33%[-b]	96.85%[-b]	96.6%[-b]	97.66%[b]	97.04%[*]	97.7%[b]	97.46%[b]	97.84%[b]	96.06%[-b]	95.53%[-b]	07 18%
Move3	96.39%[-s]	95.7%[-b]	96.82%[m]	96.21%[-b]	96.8%[s]	95.92%[-b]	96.88%[m]	96.46%[-s]	97.19%[b]	94.92%[-b]	95.03%[-b]	
Merge3	96.96%[-s]	96.41%[-b]	97.31%[m]	96.92%[-s]	97.32%[m]	96.5%[-b]	97.39%[b]	96.98%[*]	97.7%[b]	95.93%[-b]	95.9%[-b]	
Split7	96.9%[-b]	96.62%[-b]	97.14%[-m]	96.83%[-b]	97.78%[b]	97.01%[-m]	97.75%[b]	97.45%[s]	97.99%[b]	96.08%[-b]	95.85%[-b]	
Move7	96.41%[s]	95.7%[-b]	96.83%[b]	96.2%[*]	96.38%[s]	95.75%[-b]	96.71%[b]	96.17%[-s]	96.84%[b]	94.35%[-b]	94.86%[-b]	
Merge7	97.34%[b]	96.86%[-s]	97.62%[b]	97.33%[m]	97.33%[b]	96.9%[-s]	97.53%[b]	97.33%[m]	97.6%[b]	95.64%[-b]	96.05%[-b]	
Borderline20	97.17%[m]	97.1%[s]	97.26%[b]	97.2%[m]	97.35%[b]	97.16%[m]	97.27%[b]	97.28%[b]	97.42%[b]	96.58%[-m]	96.96%[*]	
Borderline100	95.7%[-s]	95.6%[-s]	95.96%[*]	95.73%[-s]	96.16%[s]	95.7%[-s]	96.11%[s]	95.97%[*]	96.3%[m]	95.23%[-b]	95.1%[-b]	
Rare20	93.15%[b]	92.95%[b]	93.27%[b]	93.05%[b]	92.93%[b]	92.61%[b]	92.94%[b]	92.73%[b]	92.99%[b]	92.1%[b]	92.59%[b]	
Rare100	71.6%[b]	69.1%[b]	72.11%[b]	67.03%[-s]	70.83%[b]	70.61%[b]	67.55%[s]	70.79%[b]	70.48%[b]	62.79%[-b]	60.35%[-b]	
					StaticI	n10				. ,		
Split3	96.06%[s]	95.84%[*]	95.91%[*]	96.4%[b]	97.3%[b]	95.64%[-s]	96.71%[b]	97.43%[b]	97.57%[b]	89.49%[-b]	96.09%[s]	95.81%
Move3	96.16%[b]	94.73%[-b]	96.26%[b]	96.29%[b]	96.39%[b]	94.69%[-b]	95.59%[m]	96.37%[b]	96.82%[b]	85.88%[-b]	95.26%[*]	
Merge3	96.79%[b]	95.61%[-s]	96.96%[b]	97.02%[b]	97.02%[b]	95.67%[*]	96.35%[b]	97.04%[b]	97.43%[b]	89.56%[-b]	96.04%[s]	
Split7	96.36%[s]	96.12%[*]	96.08%[*]	96.59%[b]	97.43%[b]	95.55%[-b]	96.54%[b]	97.42%[b]	97.62%[b]	90.1%[-b]	96.24%[*]	
Move7	96.25%[b]	94.65%[-m]	96.33%[b]	96.34%[b]	96.15%[b]	94.57%[-m]	95.68%[b]	96.29%[b]	96.62%[b]	86.53%[-b]	94.89%[-s]	95.1%
Merge7	97.21%[b]	96.16%[*]	97.36%[b]	97.47%[b]	97.13%[b]	96.18%[*]	96.76%[b]	97.43%[b]	97.52%[b]	90.4%[-b]	96.11%[*]	96.11%
Borderline20	97.21%[b]	96.98%[b]	97.12%[b]	97.32%[b]	97.2%[b]	96.85%[s]	96.69%[s]	97.37%[b]	97.37%[b]	94.08%[-b]	97.16%[b]	96.64%
Borderline100	95.73%[s]	95.45%[*]	95.64%[s]	95.86%[s]	96.02%[b]	94.69%[-b]	95.01%[-s]	96.09%[b]	96.18%[b]	90.16%[-b]	95.45%[*]	95.48%
Rare20	93.41%[b]	92.56%[b]	93.53%[b]	93.19%[b]	93.05%[b]	91.99%[m]	92.55%[b]	92.99%[b]	93.35%[b]	90.14%[-b]	92.97%[b]	91.749
Rare100	72.91%[b]	62.94%[m]	70.19%[b]	68.8%[b]	68.27%[b]	68.76%[b]	51.42%[-b]	65.76%[b]	63.83%[b]	54.11%[-b]	64.61%[b]	61.949
					StaticI	m1						
Split3	82.88%[-b]	89.88%[*]	74.12%[-b]	87.98%[-b]	58.16%[-b]	85.0%[-b]	50.0%[-b]	84.73%[-b]	82.01%[-b]	32.9%[-b]	64.2%[-b]	
Move3	84.22%[-b]	91.28%[m]	74.16%[-b]	88.87%[-m]	63.31%[-b]	83.36%[-b]	49.59%[-b]	82.48%[-b]	77.1%[-b]	29.82%[-b]	58.14%[-b]	89.59%
Merge3	88.14%[-b]	93.23%[b]	80.01%[-b]	91.68%[s]	62.83%[-b]	85.54%[-b]	53.14%[-b]	86.08%[-b]	82.98%[-b]	39.1%[-b]	64.92%[-b]	90.94%
Split7	82.88%[-b]	90.08%[*]	74.62%[-b]	87.72%[-b]	60.18%[-b]	85.78%[-b]	53.32%[-b]	85.46%[-b]	79.56%[-b]	38.07%[-b]	64.65%[-b]	90.32%
Move7	84.99%[-b]	92.21%[b]	74.64%[-b]	89.17%[-s]	62.19%[-b]	82.4%[-b]	46.07%[-b]	83.93%[-b]	75.46%[-b]	29.49%[-b]	57.16%[-b]	89.67%
Merge7	89.39%[-b]	94.12%[b]	82.11%[-b]	92.6%[m]	67.24%[-b]	85.9%[-b]	62.51%[-b]	87.98%[-b]	84.47%[-b]	45.9%[-b]	68.19%[-b]	91.19%
Borderline20	92.55%[-b]	96.34%[b]	89.44%[-b]	95.19%[-s]	64.7%[-b]	91.6%[-b]	68.54%[-b]	92.32%[-b]	89.66%[-b]	65.91%[-b]	75.55%[-b]	95.52%
Borderline100	89.3%[-b]	93.52%[*]	84.25%[-b]	93.19%[*]	55.61%[-b]	84.27%[-b]	43.86%[-b]	80.26%[-b]	67.17%[-b]	37.18%[-b]	57.48%[-b]	93.29%
Rare20	90.37%[-b]	91.24%[-b]	87.27%[-b]	92.27%[-s]				91.46%[-b]	92.95%[s]	64.7%[-b]	69.56%[-b]	
Rare100	68.79%[b]	58.12%[-b]	61.17%[s]		37.73%[-b]		42.28%[-b]	62.02%[m]	71.92%[b]	20.96%[-b]	31.26%[-b]	60.8%
					mbalance R							
StaticIm10_Im1	98.11%[b]	96.27%[-s]	97.98%[b]	98.06%[b]	98.05%[b]	95.8%[-s]	97.51%[b]	98.06%[b]	98.01%[b]		L J	
StaticIm1_Im10	93.08%[-b]	95.51%[s]	89.22%[-b]	95.78%[m]	78.05%[-b]	89.72%[-b]	73.33%[-b]	91.38%[-b]	85.05%[-b]	65.09%[-b]	81.84%[-b]	94.86%

Table 3: 30 Runs Average G-Mean on Five Dimensional Artificial Data Streams (A12 SMOClust vs Others)

Im1	97.98%[b]	97.88%[b]	98.25%[b]	97.75%[b]	92.58%[*]	93.59%[-b]	88.8%[*]	95.15%[s]	92.31%[-m]	74.32%[-b]	77.51%[-b]	96.97%
StaticIm1_Im50	93.27%[-b]	96.39%[b]	89.72%[-b]	95.78%[s]	79.14%[-b]	90.66%[-b]	75.44%[-b]	91.47%[-b]	85.77%[-b]	69.82%[-b]	82.33%[-b]	95.37%
					Double Fa	actors						
Im1+Rare100	51.35%[-b]	68.42%[b]	51.5%[-b]	46.6%[-b]	49.82%[-b]	65.26%[b]	45.57%[-b]	48.76%[-b]	52.74%[-b]	46.48%[-b]	43.65%[-b]	53.74%
Im10+Rare60	79.26%[b]	77.67%[-b]	79.8%[b]	78.3%[-b]	79.05%[b]	77.34%[-b]	79.04%[b]	78.13%[-b]	79.67%[b]	76.01%[-b]		
Split5+Im10	96.58%[-m]	95.55%[-b]	96.66%[-m]	96.74%[-s]	97.58%[b]	96.17%[-b]	97.47%[b]	97.52%[b]	97.89%[b]	92.7%[-b]	95.8%[-b]	96.94%
Im1+Borderline100	93.93%[-b]	94.87%[-s]	93.1%[-b]	94.58%[-m]	95.4%[*]	92.61%[-b]	93.39%[-m]	94.87%[*]	88.1%[*]	69.94%[-b]	73.85%[-b]	95.28%
Im10+Borderline20	97.19%[b]	96.88%[s]	97.17%[b]	97.23%[b]	97.27%[b]	96.96%[s]	97.1%[m]	97.25%[b]	97.37%[b]	96.16%[-b]	96.98%[s]	96.77%
					Complex 1	Factor						
StaticIm10_Split5	61.46%[b]	48.75%[-b]	58.21%[b]	56.93%[*]	57.75%[b]	58.28%[*]	46.11%[-b]	54.66%[b]	50.65%[b]	41.14%[-b]	40.6%[b]	56 04%
+Im1+Rare100	01.4070[D]	40.7570[-0]	J8.2170[D]	50.9570[]	51.1576[D]	J8.2070[]	40.11/0[-0]	54.0070[-D]	09.0070[D]	41.1470[-D]	40.070[-0]	50.9470
StaticIm10_Split5	84.66%[-b]	87.89%[*]	70.72%[b]	86.88%[-m]	81.07%[b]	84.9%[-s]	54.97%[b]	84 50% [b]	82.97%[-b]	51.06%[b]	59 990% [b]	97 5607
+Im1+Borderline100	04.0070[-D]	81.8370[]	19.1370[-0]	80.8870[-11]	01.9770[-D]	04.370[-5]	04.0170[-0]	04.0970[-D]	82.9170[-0]	51.0070[-D]	58.2870[-b]	81.5070
Split5+Im10	80.21%[b]	77.88%[-b]	80.43%[b]	79.4%[-m]	80.71%[b]	79.91%[s]	79.88%[b]	80.55%[b]	80.96%[b]	66.85%[-b]	78.00%[b]	70 72%
+Borderline40+Rare40	80.2170[D]	11.0070[-D]	80.4370[D]	79.470[-111]	80.7170[D]	79.9170[S]	79.8870[D]	80.3370[D]	80.9070[D]	00.0370[-D]	78.2270[-D]	19.1270
Split5+Im10	90.84%[-b]	90.06%[-b]	91.11%[-s]	90.86%[-b]	01.20%[]]	00.17%[b]	90.58%[-m]	01 250% [*]	91.4%[*]	80.55%[-b]	89.4%[-b]	01 450%
+Borderline80	90.84%[-0]	90.00%[-b]	91.11%[-s]	90.80%[-b]	91.5270[-S]	90.17%[-b]	90.58%[-m]	91.55%[']	91.470[']	80.55%[-b]	89.470[-D]	91.4370
Im10+Borderline20	91.01%[b]	90.46%[*]	91.06%[b]	90.92%[b]	91.24%[b]	90.99%[b]	91.33%[b]	91.17%[b]	91.51%[b]	89.39%[-b]	90.72%[s]	00 = 107
+Rare20	91.01%[D]	90.40%[']	91.00%[D]	90.9276[D]	91.2470[D]	90.99%[b]	91.55%[D]	91.17%[D]	91.51%[D]	89.39%[-D]	90.7276[S]	90.31%

- Based on the average G-Mean, cells are highlighted in lime / light grey when SMOClust performed better than the corresponding approach and cells are highlighted in orange / dark grey cells when SMOClust performed worse than the corresponding approach. The colour saturation scales with the absolute difference in average G-Mean between the SMOClust and the approach of the column and the saturation reaches the maximum when such difference is $\geq 10\%$. - Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against SMOClust. Presence/absence of the sign "-" in the effect size mass that the corresponding approach was worse/better than SMOClust.

Table 4: 30 Runs Average G-Mean on Five Dimensional Severely Class Imbalanced Artificial Data Streams (A12 SMOClustvs Others)

Stream	OOB	UOB	oOS	oUnder- OverB	OOB_{d}	UOB _d	$\mathrm{oOS}_{\mathrm{d}}$	oUnder- OverB _d	SMO- Gau- Noise	VFC- SMOTE	SMOTE- OB	SMO- Clust	
					St	aticIm5						L	
Split3	95.24%[*]	95.54%[s]	94.47%[-b]	96.1%[b]	96.2%[b]	95.26%[*]	94.09%[-s]	97.11%[b]	97.11%[b]	79.24%[-b]	94.26%[-b]	95.19%	
Move3	95.26%[b]	93.64%[-m]	94.74%[s]	95.93%[b]	95.46%[b]	93.48%[-b]	88.6%[-b]	95.88%[b]	96.11%[b]	71.15%[-b]	93.28%[-b]	94.23%	
Merge3	96.01%[b]	94.64%[*]	95.88%[b]	96.67%[b]	96.18%[b]	94.59%[*]	89.64%[-b]	96.65%[b]	96.87%[b]	77.97%[-b]	94.5%[*]	94.7%	
Split7	95.07%[*]	95.23%[s]	94.05%[-b]	95.82%[b]	95.46%[m]	94.61%[-s]	92.68%[-s]	96.94%[b]	96.86%[b]	77.71%[-b]	94.24%[-m]	94.99%	
Move7	95.56%[b]	93.37%[-b]	94.83%[m]	95.93%[b]	95.12%[b]	93.02%[-b]	90.63%[-b]	95.74%[b]	95.93%[b]	72.27%[-b]	92.77%[-b]	94.31%	
Merge7	96.72%[b]	95.4%[s]	96.42%[b]	97.16%[b]	96.36%[b]	95.22%[*]	92.4%[-b]	97.0%[b]	97.15%[b]	80.6%[-b]	94.37%[-b]	95.21%	
Borderline20	97.1%[b]	96.92%[s]	96.76%[*]	97.42%[b]	96.39%[-s]	96.73%[*]	94.59%[-b]	97.26%[b]	97.25%[b]	90.11%[-b]	96.87%[s]	96.69%	
Borderline100	95.63%[s]	94.99%[-s]	95.12%[-s]	95.99%[m]	95.18%[-s]	94.31%[-b]	86.67%[-b]	96.15%[b]	96.02%[m]	78.78%[-b]	95.13%[-s]	95.45%	
Rare20	93.68%[b]	92.29%[s]	93.51%[b]	93.46%[b]	92.73%[b]	91.55%[-b]	90.99%[-b]	93.04%[b]	93.66%[b]	86.71%[-b]	92.82%[b]	92.15%	
Rare100	73.34%[b]	59.64%[-s]	68.17%[b]	69.93%[b]	66.84%[b]	66.64%[b]	47.73%[-b]	61.41%[s]	60.23%[*]	44.7%[-b]	61.16%[s]	60.53%	
StaticIm3													
Split3	93.67%[-b]	95.28%[m]	91.81%[-b]	95.12%[m]	92.24%[-b]	94.84%[s]	87.39%[-b]	96.2%[b]	95.94%[b]	67.27%[-b]	90.19%[-b]	94.51%	
Move3	93.84%[s]	93.04%[-s]	91.87%[-b]	95.21%[b]	93.03%[*]	92.35%[-m]	80.77%[-b]	94.86%[b]	94.83%[b]	58.53%[-b]	88.14%[-b]	93.27%	
Merge3	95.04%[b]	94.03%[*]	93.86%[*]	96.19%[b]	93.78%[*]	93.55%[*]	84.34%[-b]	95.86%[b]	96.09%[b]	67.23%[-b]	91.21%[-b]	93.7%	
Split7	93.2%[-b]	95.05%[b]	90.86%[-b]	94.76%[m]	90.64%[-b]	93.94%[*]	84.29%[-b]	95.8%[b]	95.76%[b]	67.82%[-b]	and the second state	94.15%	
Move7	94.33%[b]	92.86%[-m]	92.31%[-b]	95.33%[b]	92.73%[-s]	92.23%[-b]	82.03%[-b]	94.74%[b]	94.4%[b]	59.0%[-b]	86.51%[-b]		
Merge7	95.93%[b]	94.97%[*]	94.81%[*]	96.79%[b]	94.35%[*]	94.61%[-s]	85.91%[-b]	96.23%[b]	95.85%[b]	71.88%[-b]	90.64%[-b]		
Borderline20	96.58%[*]	96.8%[s]	95.97%[-b]	97.23%[b]	92.88%[-b]	96.43%[-s]	91.48%[-b]	96.83%[s]	96.53%[*]	86.63%[-b]		96.66%	
Borderline100	94.97%[-s]	94.59%[-m]	93.93%[-b]	95.77%[m]	91.38%[-b]	93.86%[-b]	70.99%[-b]	95.81%[m]	94.69%[*]	65.16%[-b]	92.37%[-b]	95.2%	
Rare20	93.43%[b]	92.09%[-m]	92.97%[b]	93.59%[b]	90.17%[-b]	90.87%[-b]	90.02%[-b]	93.33%[b]	94.44%[b]	84.19%[-b]	91.73%[-b]		
Rare100	71.72%[b]	62.22%[b]	66.12%[b]	69.63%[b]	64.93%[b]	62.23%[b]	53.0%[-b]	62.63%[b]	61.83%[s]	35.52%[-b]	43.99%[-b]	59.92%	
						aticIm1							
Split3	84.29%[-b]	90.98%[*]	75.65%[-b]	88.67%[-b]	61.71%[-b]	85.62%[-b]	54.23%[-b]	85.79%[-b]	82.17%[-b]	36.52%[-b]		90.91%	
Move3	82.76%[-b]	91.14%[m]	71.59%[-b]	87.92%[-b]	70.48%[-b]	85.17%[-b]	46.61%[-b]	81.17%[-b]	76.86%[-b]	26.69%[-b]		89.55%	
Merge3	86.61%[-b]	92.27%[b]	77.68%[-b]	90.9%[s]	67.28%[-b]	85.85%[-m]	52.57%[-b]	84.94%[-b]	82.54%[-b]	36.97%[-b]	62.23%[-b]	90.22%	
Split7	82.03%[-b]	89.38%[*]	73.04%[-b]	87.2%[-b]	58.85%[-b]	85.31%[-b]	52.58%[-b]	84.55%[-b]	79.12%[-b]	38.28%[-b]	65.42%[-b]	90.25%	
Move7	83.3%[-b]	91.7%[b]	73.53%[-b]	88.48%[-m]	69.71%[-b]	81.62%[-b]	47.2%[-b]	81.25%[-b]	76.87%[-b]	28.27%[-b]		89.36%	
Merge7	88.16%[-b]	93.49%[b]	81.61%[-b]	92.15%[m]	74.91%[-b]	85.81%[-b]	60.94%[-b]	86.48%[-b]	84.54%[-b]	45.07%[-b]		91.15%	
Borderline20	92.42%[-b]	96.49%[m]	89.57%[-b]	95.32%[-m]	72.26%[-b]	91.37%[-b]	70.68%[-b]	92.25%[-b]	90.22%[-b]	64.94%[-b]	76.87%[-Ь]	95.9%	
Borderline100	89.38%[-b]	93.84%[*]	84.72%[-b]	93.39%[-s]	55.97%[-b]	86.42%[-b]	43.44%[-b]	77.92%[-b]	67.67%[-b]	38.66%[-b]	61.04%[-b]	93.66%	
Rare20	90.09%[-b]	91.35%[-b]	87.3%[-b]	92.24%[-m]	71.2%[-b]	80.35%[-b]	70.04%[-b]	90.97%[-b]	92.31%[*]	63.8%[-b]		92.89%	
Rare100	68.79%[b]	58.9%[-b]	61.54%[*]	67.6%[b]	43.52%[-b]	45.74%[-b]	42.37%[-b]	62.08%[*]	71.26%[b]	21.13%[-b]	31.55%[-b]	61.5%	
					Sta	aticIm07							

14

SMOClust

Split3	77.51%[-b]	89.86%[*]	65.19%[-b]	83.34%[-b]	45.48%[-b]	80.25%[-b]	36.62%[-b]	74.97%[-b]	70.32%[-b]	22.27%[-b]	44.18%[-b]	89.45%
Move3	74.0%[-b]	90.55%[b]	57.69%[-b]	81.13%[-b]	43.83%[-b]	76.82%[-b]	34.75%[-b]	70.38%[-b]	63.8%[-b]	16.96%[-b]	36.23%[-b]	87.26%
Merge3	79.16%[-b]	91.48%[b]	66.0%[-b]	85.44%[-s]	47.07%[-b]	77.02%[-b]	37.31%[-b]	75.7%[-b]	70.53%[-b]	22.74%[-b]	41.93%[-b]	87.18%
Split7	75.59%[-b]	88.17%[*]	63.64%[-b]	81.65%[-b]	41.83%[-b]	77.0%[-b]	37.01%[-b]	74.84%[-b]	70.73%[-b]	25.2%[-b]	45.21%[-b]	88.22%
Move7	74.2%[-b]	91.4%[b]	59.56%[-b]	82.08%[-b]	45.21%[-b]	76.38%[-b]	36.29%[-b]	69.28%[-b]	62.5%[-b]	18.41%[-b]	36.33%[-b]	87.43%
Merge7	81.7%[-b]	92.91%[b]	71.34%[-b]	87.48%[-m]	50.87%[-b]	81.99%[-b]	48.24%[-b]	78.85%[-b]	75.74%[-b]	35.29%[-b]	49.15%[-b]	88.97%
Borderline20	89.33%[-b]	96.39%[b]	84.19%[-b]	93.29%[-b]	56.19%[-b]	88.59%[-b]	48.37%[-b]	89.35%[-b]	80.2%[-b]	53.25%[-b]	59.4%[-b]	95.1%
Borderline100	84.99%[-b]	93.42%[s]	77.45%[-b]	90.82%[-b]	39.67%[-b]	76.94%[-b]	30.87%[-b]	66.34%[-b]	56.31%[-b]	27.88%[-b]	42.17%[-b]	92.59%
Rare20	87.34%[-b]	90.98%[-b]	82.05%[-b]	90.52%[-b]	51.72%[-b]	71.26%[-b]	45.94%[-b]	87.73%[-b]	87.15%[-b]	52.42%[-b]	54.21%[-b]	92.51%
Rare100	68.59%[b]	58.39%[-b]	58.71%[-b]	66.68%[b]	25.99%[-Ь]	40.18%[-b]	28.49%[-b]	62.65%[m]	70.08%[b]	15.94%[-b]	24.29%[-b]	61.97%
					St	aticIm05						
Split3	67.41%[-b]	89.06%[m]	51.8%[-b]	75.26%[-b]	28.85%[-b]	70.03%[-b]	24.15%[-b]	61.4%[-b]	58.01%[-b]	11.44%[-b]	24.19%[-b]	86.92%
Move3	60.73%[-b]	89.92%[b]	42.99%[-b]	71.27%[-b]	25.71%[-b]	60.37%[-b]	22.91%[-b]	52.3%[-b]	49.96%[-b]	7.7%[-b]	19.35%[-b]	82.97%
Merge3	68.16%[-b]	90.63%[b]	52.0%[-b]	77.04%[-b]	26.29%[-b]	63.85%[-b]	25.45%[-b]	60.63%[-b]	53.63%[-b]	11.47%[-b]	23.17%[-b]	84.28%
Split7	66.73%[-b]	87.47%[s]	52.25%[-b]	74.33%[-b]	29.97%[-b]	64.71%[-b]	26.36%[-b]	64.25%[-b]	59.18%[-b]	17.2%[-b]	26.46%[-b]	85.98%
Move7	61.93%[-b]	91.08%[b]	44.91%[-b]	72.82%[-b]	32.82%[-b]	57.0%[-b]	25.03%[-b]	54.0%[-b]	49.63%[-b]	10.96%[-b]	19.4%[-b]	84.18%
Merge7	72.74%[-b]	92.48%[b]	59.63%[-b]	80.73%[-b]	39.25%[-b]	69.5%[-b]	31.25%[-b]	70.66%[-b]	64.57%[-b]	22.95%[-b]	30.44%[-b]	86.78%
Borderline20	84.19%[-b]	96.29%[b]	76.48%[-b]	90.35%[-b]	30.95%[-b]	79.13%[-b]	36.96%[-b]	77.36%[-b]	69.86%[-b]	40.47%[-b]	40.21%[-b]	93.92%
Borderline100	77.54%[-b]	93.15%[b]	68.43%[-b]	87.03%[-b]	26.19%[-b]	57.71%[-b]	25.39%[-b]	52.65%[-b]	47.62%[-b]	15.48%[-b]	23.16%[-b]	91.05%
Rare20	82.38%[-b]	90.85%[-m]	74.39%[-b]	87.71%[-b]	33.53%[-b]	60.14%[-b]	38.56%[-b]	81.84%[-b]	78.82%[-b]	39.53%[-b]	35.61%[-b]	91.51%
Rare100	65.62%[m]	59.16%[-b]	55.02%[-b]	65.53%[b]	14.14%[-b]	32.84%[-b]	26.25%[-b]	56.47%[-b]	67.64%[m]	9.8%[-b]	15.85%[-b]	63.7%
					St	aticIm03						
Split3	49.77%[-b]	88.35%[b]	33.95%[-b]	58.57%[-b]	14.32%[-b]	45.41%[-b]	13.85%[-b]	42.25%[-b]	39.95%[-b]	1.65%[-b]	7.15%[-b]	81.31%
Move3	40.57%[-b]	88.81%[b]	22.43%[-b]	49.93%[-b]	18.07%[-b]	40.61%[-b]	10.55%[-b]	35.38%[-b]	35.13%[-b]	0.28%[-Ь]	5.35%[-b]	75.16%
Merge3	49.77%[-b]	89.5%[b]	31.09%[-b]	59.18%[-b]	17.83%[-b]	46.84%[-b]	14.48%[-b]	41.79%[-b]	35.16%[-b]	1.09%[-b]	6.67%[-b]	76.22%
Split7	50.45%[-b]	86.36%[b]	32.87%[-b]	57.4%[-b]	15.86%[-b]	48.09%[-b]	14.5%[-b]	41.15%[-b]	39.14%[-b]	2.25%[-b]	8.08%[-b]	80.4%
Move7	41.14%[-b]	90.27%[b]	23.33%[-b]	53.4%[-b]	18.43%[-b]	33.73%[-b]	10.78%[-b]	33.98%[-b]	34.82%[-b]	0.87%[-b]	5.4%[-b]	75.72%
Merge7	56.06%[-b]	91.5%[b]	38.89%[-b]	65.82%[-b]	23.33%[-b]	52.59%[-b]	16.89%[-b]	52.72%[-b]	49.98%[-b]	4.99%[-b]	10.22%[-b]	78.12%
Borderline20	70.16%[-b]	96.09%[b]	55.01%[-b]	81.64%[-b]	23.08%[-b]	58.81%[-b]	19.34%[-b]	56.61%[-b]	43.26%[-b]	16.27%[-b]	14.33%[-b]	91.8%
Borderline100	61.02%[-b]	92.72%[b]	49.14%[-b]	76.31%[-b]	18.34%[-b]	43.35%[-b]	15.68%[-b]	37.21%[-b]	37.93%[-b]	2.46%[-b]	6.6%[-b]	88.13%
Rare20	69.26%[-b]	90.63%[s]	53.43%[-b]	79.4%[-b]	19.93%[-Ь]	44.64%[-b]	19.35%[-b]	59.5%[-b]	53.38%[-b]	16.35%[-b]	12.85%[-b]	89.69%
Rare100	56.62%[-b]	58.64%[-b]	45.25%[-b]	60.56%[-b]	10.95%[-b]	24.78%[-b]	16.73%[-b]	47.49%[-b]	57.16%[-b]	5.22%[-b]	7.17%[-b]	64.62%

- Based on the average G-Mean, cells are highlighted in lime / light grey when SMOClust performed better than the corresponding approach and cells are highlighted in orange / dark grey cells when SMOClust performed worse than the corresponding approach. The colour saturation scales with the absolute difference in average G-Mean between the SMOClust and the approach of the column and the saturation reaches the maximum when such difference is $\geq 10\%$. - Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against SMOClust. Presence/absence of the sign "-" in the effect size mass that the corresponding approach was worse/better than SMOClust.

	L	1	0				`			/		
Stream	OOB	UOB	oOS	oUnder- OverB	OOB_d	UOB_d	$\mathrm{oOS}_{\mathrm{d}}$	oUnder- OverB _d	SMO- Gau- Noise	VFC- SMOTE	SMOTE- OB	SMO- Clust
	1				StaticIr	m30						
Split3	96.93%[-b]	96.76%[-b]	97.08%[-b]	97.06%[-b]	98.19%[*]	97.31%[-b]	98.61%[b]	97.85%[-b]	98.53%[b]	89.56%[-b]	97.07%[-Ъ]	08 23%
Move3	95.46%[-b]	94.33%[-b]	96.15%[-b]	95.22%[-b]	97.28%[-s]	94.87%[-b]	98.09%[b]	96.43%[-b]	97.89%[b]	83.22%[-b]	95.14%[-b]	
Merge3	96.99%[-b]	94.33%[-b] 95.98%[-b]	97.69%[-b]	96.71%[-b]	97.28%[-8] 98.45%[-b]	96.87%[-b]	98.85%[b]	90.43%[-b] 97.73%[-b]	97.89%[b] 98.82%[b]	92.15%[-b]	97.66%[-b]	
Split7	96.07%[-b]	95.33%[-b]	96.36%[-b]	95.87%[-b]	97.14%[*]	93.9%[-b]	97.84%[b]	96.0%[-b]	98.82 %[b] 97.63%[b]	88.05%[-b]	97.00%[-b] 95.16%[-b]	
Move7	93.6%[-b]	91.96%[-b]	94.45%[-b]	92.88%[-b]	94.95%[-s]	90.97%[-b]	96.57%[b]	92.99%[-b]	96.29%[b]	81.39%[-b]		
Merge7	96.12%[-b]	95.24%[-b]	96.64%[-b]	96.04%[-b]	96.94%[-b]	95.32%[-b]	98.27%[b]	96.11%[-b]	98.2%[b]	90.96%[-b]		
Borderline20	99.47%[-m]	99.45%[-m]	90.04%[-b] 99.48%[-s]	99.47%[-b]	99.52%[s]	99.48%[-s]	99.53%[s]	99.51%[*]	99.53%[s]	98.3%[-b]	99.46%[-m]	
Borderline100	99.34%[-h]		99.48%[-b]	99.35%[-b]	99.32%[s] 99.39%[-s]	99.32%[-b]	99.33%[8] 99.42%[*]	99.36%[-m]	99.33%[s] 99.49%[b]	97.92%[-b]		
Rare20	93.79%[b]	99.33%[-D] 93.75%[b]	93.82%[b]	93.77%[b]	93.66%[b]	99.327%[-D] 93.43%[*]	93.72%[b]	93.51%[b]	99.49%[b] 93.68%[b]		99.33%[-b] 93.34%[-m]	
Rare100	68.44%[-b]	67.27%[-b]	93.82 %[D] 68.7%[-b]		67.72%[-b]		67.85%[-b]	67.69%[-b]	67.8%[-b]	62.38%[-b]	56.09%[-h]	
Rafe100	08.4470[-D]	01.2170[-D]	08.770[-D]	02.0970[-0]	StaticIr		07.8576[-D]	07.0970[-D]	07.870[-D]	02.3870[-D]		08.957
Split3	96.69%[-b]	95.74%[-b]	96.75%[-b]	96.96%[-b]	97.98%[s]	94.08%[-b]	98.27%[b]	97.98%[s]	98.21%[b]	78.57%[-b]	97.39%[-b]	07 000
		L J										
Move3	95.52%[-b]	92.32%[-b] 94.71%[-b]	95.86%[-b] 97.51%[-b]	95.52%[-b]	97.24%[b]	90.83%[-b]	97.71%[b]	96.85%[s]	97.55%[b]	69.92%[-b] 83.46%[-b]		
Merge3	97.21%[-b]			97.05%[-b]	98.52%[s]	94.86%[-b]	98.68%[b]	98.11%[-b]	98.69%[b]		97.77%[-b]	
Split7 Move7			95.87%[-b]	95.92%[-b]	97.02%[m]	87.8%[-b]	97.35%[b]	96.61%[*]	97.26%[b]	78.85%[-b] 68.29%[-b]	95.31%[-b]	
	93.61%[-b] 96.2%[-b]		94.1%[-b]	93.36%[-b] 96.36%[-b]	95.32%[b]	86.08%[-b] 93.64%[-b]	96.03%[b] 98.05%[b]	93.92%[-b]	95.9%[b] 98.02%[b]	82.56%[-b]	92.37%[-b] 96.64%[-b]	
Merge7 Borderline20		93.82%[-b]	96.54%[-b]		97.52%[s]		98.05%[b] 99.43%[-b]	96.28%[-b]	98.02%[b] 99.47%[-s]			
Borderline100	99.44%[-b] 99.29%[-b]	99.42%[-b] 99.3%[-b]	99.41%[-b]	99.47%[-m]		99.37%[-b]		99.48%[-s]		94.16%[-b] 92.51%[-b]	99.43%[-b] 99.3%[-m]	
Rare20	99.29%[-b] 93.85%[b]	99.5%[-b] 93.67%[-b]	99.27%[-b] 93.92%[b]	99.33%[-s] 93.79%[b]	99.3%[-m] 93.67%[-b]	99.14%[-b] 92.81%[-b]	99.33%[-s] 93.74%[*]	99.33%[-s] 93.67%[-b]	99.44 %[m] 93.79%[b]		99.3%[-m] 93.73%[-m]	
Rare100	68.69%[-b]		67.54%[-b]									
nare100	08.09%[-D]	01.0%[-D]	07.34%[-D]	04.36%[-D]	04.95%[-D] StaticI	66.67%[-b]	04.06%[-D]	03.10%[-D]	04.3770[-D]	-05.07%[-D]	57.31%[-b]	09.1%
C., 114.9	00.007[1]	04.0007[1]	01 1707[1]	0.2 C107 [1]			00.0207[1]	00.0507[1]	00.0407[1]	10.0907[1]	70.907[1]	00.250
Split3	92.62%[-b]	84.82%[-b]	91.17%[-b]	93.61%[-b]		70.84%[-b]			92.94%[-b]			96.35%
Move3	91.09%[-b]	81.03%[-b]	89.5%[-b]	92.16%[-b]	68.97%[-b]	64.54%[-b]	87.83%[-b]	78.87%[-b]	91.72%[-b]	a ser e construction de la const	70.61%[-b]	95.61%
Merge3	94.54%[-b]	90.06%[-b]	93.79%[-b]	95.41%[-b]	77.41%[-b]		93.25%[-b]	88.34%[-b]	95.42%[-b]		85.27%[-b]	97.93%
${ m Split7} { m Move7}$	89.73%[-b] 87.09%[-b]	75.71%[-b]	87.88%[-b]	91.15%[-b]	63.24%[-b]	63.2%[-b]	87.41%[-b]	76.54%[-b]	90.78%[-b]		59.28%[-b]	94.877
		73.68%[-b]	84.53%[-b]	88.9%[-b]		55.73%[-b]	82.0%[-b]	62.05%[-b]	87.95%[-b]			92.92%
Merge7	93.08%[-b]	88.23%[-b]	91.87%[-b]	94.47%[-b]	73.99%[-b]		90.75%[-b]		93.94%[-b]		76.63%[-b]	
Borderline20	98.19%[-b]	99.31%[-s]	97.97%[-b]		81.89%[-b]		96.83%[-b]		97.92%[-b]		97.75%[-b]	
Borderline100 Rare20	97.95%[-b] 93.69%[-b]	99.0%[-b]	97.61%[-b]	98.68%[-b]	82.35%[-b]		97.1%[-b]	98.15%[-b]	98.31%[-b]		97.34%[-b]	
		92.78%[-b]	93.77%[-b]	94.12%[-b]		79.71%[-b]			94.22%[-m]			
Rare100	62.82%[-b]	59.84%[-b]	61.6%[-b]		36.31%[-b]		55.35%[-b]	52.08%[-b]	56.94%[-b]	17.43%[-b]	38.34%[-b]	(2.51)
				11	mbalance Ra	atio Drift						
QL 11 10 1 1	00 0107 01	00.0007[1]	00 5407 [1]	00.0507.11	00 000/51	00 1007 [1]	00 5407 51	00 050711	00 00/11	40 1007 [1]	00 5 507 51	00 450
StaticIm10_Im1 StaticIm1_Im10	99.71%[b] 98.96%[-b]	99.22%[-b] 98.77%[-b]	99.74%[b] 98.67%[-b]	99.65%[b]		99.19%[-b]	99.74%[b]	99.65%[b] 98.92%[-b]		42.16%[-b]	99.57%[b] 96.94%[-b]	

Table 5: [2D] 30 Runs Average G-Mean on Artificial Data Streams (A12 SMOClust vs Others)

Im1	99.65%[b]	99.56%[b]	99.77%[b]	99.56%[b]	99.65%[b]	98.85%[*]	99.77%[b]	99.53%[b]	99.71%[b]	49.5%[-b]	99.54%[b]	99.44%
StaticIm1_Im50	98.98%[-b]	99.01%[-b]	98.7%[-b]	99.28%[-m]			98.5%[-b]	98.92%[-b]	98.84%[-b]	62.88%[-b]	97.01%[-b]	99.4%
					Double Fa	actors						
Im1+Rare100	48.84%[-b]	67.45%[b]	50.07%[-b]	44.91%[-b]	48.1%[-b]	58.6%[m]	48.67%[-b]	48.17%[-b]	53.62%[-s]	41.04%[-b]	43.28%[-b]	53.73%
Im10+Rare60	78.57%[*]	78.11%[-b]	78.97%[b]	78.2%[-b]	78.28%[-b]	77.63%[-b]	78.84%[b]	77.85%[-b]	78.93%[b]	71.36%[-b]	79.99%[b]	78.62%
Split5+Im10	96.97%[-b]	94.44%[-b]	96.64%[-b]	96.99%[-b]	97.79%[b]	92.14%[-b]	98.13%[b]	97.11%[-m]	98.03%[b]	84.45%[-b]	96.7%[-b]	97.5%
Im1+Borderline100	99.17%[-b]	99.3%[-b]	99.08%[-b]	99.22%[-b]	98.16%[-b]	99.23%[-m]	99.13%[-b]	99.21%[-b]	99.26%[*]	50.19%[-b]	99.21%[-b]	99.4%
Im10+Borderline20	99.45%[-m]	99.37%[-b]	99.43%[-m]	99.47%[-s]	99.47%[-s]	99.46%[-s]	99.5%[*]	99.48%[-s]	99.52%[*]	96.29%[-b]	99.42%[-b]	99.5%
					Complex 1	Factor						
StaticIm10_Split5	55.24%[-b]	47.75%[-b]	54.77%[-b]	53.1%[-b]	52.7%[-b]	EE 7207[]	ED 9607[1]	52.52%[-b]	60.56%[*]	25 2607 [1.]	40.64%[-b]	CO 4507
+Im1+Rare100	JJ.2470[-D]	41.1570[-D]	54.7770[-D]	55.170[-b]	52.770[-D]	55.7570[-S]	52.8070[-D]	52.5270[-D]	00.3070[]	55.2070[-D]	40.0470[-0]	00.4570
StaticIm10_Split5	85.08%[b]	85.79%[b]	84.35%[b]	84.72%[b]	76.20%[b]	77.87%[m]	75.2%[b]	77.03%[-m]	76 50% [m]	40 54% [b]	50.20% [b]	70 90%
+Im1+Borderline100	63.0670[D]	83.7970[D]	64.5570[D]	04.7270[D]	70.3270[-D]	//.8//o[iii]	75.570[-D]	77.0370[-111]	70.5970[-111]	40.3470[-0]	30.370[-b]	10.370
Split5+Im10	79.65%[b]	79.24%[s]	80.43%[b]	79.22%[s]	78.72%[-s]	78.43%[-s]	79.04%[*]	78.95%[*]	79.08%[s]	70.01%[-b]	79.25%[s]	78.9%
+Borderline40+Rare40	79.0376[D]	79.2470[S]	80.4370[D]	79.2270[S]	10.1270[-8]	78.4370[-8]	79.0470[]	18.9570[]	79.0670[S]	70.0176[-D]	79.2070[S]	10.970
Split5+Im10	89.55%[b]	89.74%[b]	90.28%[b]	89.1%[b]	86.98%[*]	85.78%[-m]	97 0607 [*]	87.03%[*]	87.1%[*]	77.49%[-b]	88.12%[b]	97 0707
+Borderline80	89.55%[D]	69.7470[D]	90.28%[D]	89.1%[D]	80.9870[1]	85.78%[-III]	87.00%[']	87.03%[']	87.170[1]	//.49%[-D]	00.1270[D]	81.0170
Im10+Borderline20	93.23%[-m]	02.0607[1]	93.25%[-s]	93.2%[-b]	93.29%[*]	02 1007 []	02.2407[]	02.9207[]	02.2607[1.]	01.1707[h]	93.21%[-b]	02.007
+Rare20	95.25%[-m]	92.90%[-D]	93.23%[-S]	93.2%[-D]	95.29%[*]	93.19%[-b]	95.54%[m]	93.23%[-m]	93.30%[D]	91.17%[-D]	93.21%[-D]	95.29%

- Based on the average G-Mean, cells are highlighted in lime / light grey when SMOClust performed better than the corresponding approach and cells are highlighted in orange / dark grey cells when SMOClust performed worse than the corresponding approach. The colour saturation scales with the absolute difference in average G-Mean between the SMOClust and the approach of the column and the saturation reaches the maximum when such difference is $\geq 10\%$. - Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against SMOClust. Presence/absence of the sign "-" in the effect size mass that the corresponding approach was worse/better than SMOClust.

					0							/
Stream	OOB	UOB	oOS	oUnder- OverB	OOB_d	UOB_d	$\mathrm{oOS}_{\mathrm{d}}$	oUnder- OverB _d	SMO- Gau- Noise	VFC- SMOTE	SMOTE- OB	SMO- Clust
StaticIm5												
Split3	96.26%[-b]	94.21%[-b]	96.23%[-b]	96.76%[-b]	96.93%[-b]	87.48%[-b]	97.56%[*]	97.81%[m]	97.74%[m]	63.78%[-b]	97.08%[-b]	97.53%
Move3	95.3%[-b]	89.73%[-b]	95.5%[-b]	95.38%[-b]	96.44%[*]	87.8%[-b]	97.13%[b]	96.55%[s]	97.12%[b]	51.18%[-b]	95.6%[-b]	96.42%
Merge3	97.1%[-b]	93.38%[-b]	97.32%[-b]	97.08%[-b]	98.07%[-s]	93.19%[-b]	98.37%[s]	97.94%[-m]	98.51%[b]	69.83%[-b]	97.6%[-b]	
Split7	95.25%[-b]	89.69%[-b]	95.21%[-b]	95.6%[-b]	95.69%[-m]	81.97%[-b]	96.57%[s]	96.21%[*]	96.68%[b]	65.42%[-b]	94.96%[-b]	96.34%
Move7	93.15%[-b]	84.98%[-b]	93.35%[-b]	93.1%[-b]	94.16%[s]	78.3%[-b]	95.19%[b]	93.54%[-s]	95.24%[b]	47.38%[-b]	91.55%[-b]	94.14%
Merge7	96.23%[-b]	93.15%[-b]	96.45%[-b]	96.43%[-b]	97.15%[-s]	93.06%[-b]	97.84%[b]	96.38%[-b]	97.86%[b]	70.53%[-b]	96.16%[-b]	97.41%
Borderline20	99.33%[-b]	99.39%[-b]	99.28%[-b]	99.44%[-b]	98.16%[-b]	99.28%[-b]	99.24%[-b]	99.41%[-b]	99.41%[-b]	89.66%[-Ь]	99.4%[-b]	
Borderline100	99.2%[-b]	99.25%[-b]	99.16%[-b]	99.29%[-m]	98.28%[-b]	98.89%[-b]	99.17%[-b]	99.25%[-b]	99.41%[m]	84.65%[-b]	99.25%[-b]	99.35%
Rare20	94.0%[*]	93.54%[-b]	94.1%[b]	93.87%[-b]	92.87%[-b]	91.79%[-b]	93.8%[-b]	93.67%[-b]	94.13%[b]	86.28%[-b]	93.77%[-b]	94.01%
Rare100	68.78%[-b]	53.22%[-b]	65.87%[-b]	65.5%[-b]	60.87%[-b]	65.21%[-b]	60.57%[-b]	57.55%[-b]	59.83%[-b]	41.95%[-b]	59.05%[-b]	69.18%
StaticIm3												
Split3	95.75%[-b]	93.0%[-b]	95.55%[-b]	96.31%[-b]	94.8%[-b]	84.39%[-b]	96.69%[-b]	96.64%[*]	97.09%[-s]	47.52%[-b]	96.41%[-b]	97.26%
Move3	94.92%[-b]	88.37%[-b]	94.69%[-b]	95.09%[-b]	94.72%[-b]	86.62%[-b]	95.96%[-s]	95.33%[-b]	96.4%[s]	29.26%[-b]	94.77%[-b]	96.23%
Merge3	96.92%[-b]	92.56%[-b]	96.9%[-b]	97.0%[-b]	97.02%[-b]	92.4%[-b]	97.76%[-b]	97.23%[-b]	98.1%[-s]	54.51%[-b]	97.19%[-b]	98.16%
Split7	94.66%[-b]	84.86%[-b]	94.08%[-b]	95.02%[-b]	91.72%[-b]	79.73%[-b]	95.24%[-b]	94.99%[-m]	95.99%[-s]	51.39%[-b]	92.71%[-b]	96.13%
Move7	92.6%[-b]	82.01%[-b]	92.2%[-b]	92.8%[-b]	90.52%[-b]	76.74%[-b]	93.49%[-m]	91.97%[-b]	94.21%[s]	28.71%[-b]	90.01%[-b]	93.97%
Merge7	95.97%[-b]	92.05%[-b]	96.05%[-b]	96.38%[-b]	95.2%[-b]	91.75%[-b]	97.04%[-m]	95.94%[-b]	97.48%[s]	59.75%[-b]	95.47%[-b]	97.39%
Borderline20	99.18%[-b]	99.36%[-b]	99.05%[-b]	99.36%[-b]	95.01%[-b]	99.22%[-b]	98.8%[-b]	99.26%[-b]	99.22%[-b]	83.12%[-b]	99.28%[-b]	99.51%
Borderline100	99.07%[-b]	99.14%[-b]	98.92%[-b]	99.23%[-b]	95.87%[-b]	98.78%[-b]	98.92%[-b]	99.15%[-b]	99.27%[-m]	66.83%[-b]	99.19%[-b]	99.37%
Rare20	94.05%[-b]	93.45%[-b]	94.14%[-b]	94.03%[-b]	90.51%[-b]	90.39%[-b]	93.69%[-b]	93.69%[-b]	94.38%[m]	77.65%[-b]	93.75%[-b]	94.29%
Rare100	67.01%[-b]	52.07%[-b]	63.98%[-b]	65.69%[-b]	56.76%[-b]	58.21%[-b]	58.45%[-b]	53.06%[-b]	57.32%[-b]	31.33%[-b]	44.84%[-b]	70.06%
StaticIm1												
Split3	92.71%[-b]	85.74%[-b]	91.24%[-b]	93.65%[-b]	76.82%[-b]	71.12%[-b]	89.85%[-b]	88.04%[-b]	93.45%[-b]	12.28%[-b]	76.59%[-b]	96.5%
Move3	91.27%[-b]	81.03%[-b]	89.76%[-b]	92.32%[-b]	68.52%[-b]	64.46%[-b]	87.73%[-b]	82.06%[-b]	91.85%[-b]	1.49%[-b]	65.49%[-b]	95.55%
Merge3	94.82%[-b]	89.41%[-b]	94.13%[-b]	95.51%[-b]	79.99%[-Ъ]	79.37%[-b]	93.15%[-b]	90.39%[-b]	95.57%[-b]	17.29%[-b]	84.32%[-b]	97.95%
Split7	89.82%[-b]	75.83%[-b]	87.8%[-b]	91.21%[-b]	61.89%[-b]	65.49%[-b]	87.41%[-b]	78.94%[-b]	91.36%[-b]	21.13%[-b]	59.59%[-b]	94.95%
Move7	87.15%[-b]	73.77%[-b]	84.62%[-b]	89.05%[-b]	60.65%[-b]	57.18%[-b]	82.42%[-b]	69.44%[-b]	87.99%[-b]	1.35%[-b]	43.97%[-b]	92.87%
Merge7	93.45%[-b]	88.61%[-b]	92.3%[-b]	94.84%[-b]	77.92%[-b]	80.93%[-b]	90.99%[-b]	86.07%[-b]	94.13%[-b]	30.3%[-b]	77.94%[-b]	96.94%
Borderline20	98.17%[-b]	99.3%[-s]	97.8%[-b]	98.75%[-b]	75.25%[-b]	98.24%[-b]	96.63%[-b]	97.85%[-b]	97.83%[-b]	44.38%[-b]	97.35%[-b]	99.36%
Borderline100	98.09%[-b]	98.96%[-b]	97.61%[-b]	98.64%[-b]	79.41%[-b]	97.63%[-b]	96.84%[-b]	98.17%[-b]	98.19%[-b]	19.4%[-b]	96.9%[-b]	99.31%
Rare20	93.68%[-b]	92.89%[-b]	93.69%[-b]	94.05%[-b]	73.67%[-b]	80.05%[-b]	92.23%[-b]	92.28%[-b]	94.57%[-s]	37.63%[-b]	90.35%[-b]	94.82%
Rare100	62.68%[-b]	57.72%[-b]	61.24%[-b]	64.07%[-b]	31.76%[-b]	39.77%[-b]	55.4%[-b]	52.83%[-b]	56.63%[-b]	17.6%[-b]	38.16%[-b]	72.29%
StaticIm07												
Split3	90.14%[-b]	82.64%[-b]	88.33%[-b]	91.65%[-b]	61.98%[-b]	56.83%[-b]	86.9%[-b]	81.45%[-b]	91.14%[-b]	6.22%[-b]	58.28%[-b]	95.84%
Move3	88.4%[-b]	79.1%[-b]	86.04%[-b]	90.04%[-b]	50.73%[-b]	43.13%[-b]	83.11%[-b]	69.94%[-b]	87.97%[-b]	0.29%[-b]	37.64%[-b]	94.64%
												-

Table 6: [2D] 30 Runs Average G-Mean on Severely Class Imbalanced Artificial Data Streams (A12 SMOClust vs Others)

Merge3	93.08%[-b]	88.02%[-b]	92.13%[-b]	94.23%[-b]	66.29%[-b]	73.48%[-b]	89.69%[-b]	81.7%[-b]	93.21%[-b]	8.39%[-b]	69.83%[-b]	97.45%
Split7	86.43%[-b]	74.12%[-b]	84.09%[-b]	88.5%[-b]	49.38%[-b]	54.93%[-b]	82.33%[-b]	68.52%[-b]	88.42%[-b]	15.23%[-b]	45.7%[-b]	93.09%
Move7	82.27%[-b]	73.03%[-b]	78.95%[-b]	85.64%[-b]	45.01%[-b]	43.87%[-b]	76.71%[-b]	50.16%[-b]	83.67%[-b]	0.11%[-b]	22.3%[-b]	90.64%
Merge7	90.95%[-b]	87.27%[-b]	89.29%[-b]	93.12%[-b]	64.4%[-b]	75.23%[-b]	87.39%[-b]	74.94%[-b]	91.68%[-b]	22.0%[-b]	65.47%[-b]	96.06%
Borderline20	97.42%[-b]	99.23%[m]	96.97%[-b]	98.28%[-b]	76.78%[-b]	97.06%[-b]	94.82%[-b]	96.54%[-b]	96.71%[-b]	27.63%[-b]	94.04%[-b]	99.12%
Borderline100	97.31%[-b]	98.86%[-b]	96.57%[-b]	98.14%[-b]	76.49%[-b]	95.77%[-b]	94.28%[-b]	96.57%[-b]	97.1%[-b]	12.63%[-b]	91.56%[-b]	99.16%
Rare20	93.36%[-b]	92.53%[-b]	93.08%[-b]	93.81%[-b]	64.74%[-b]	76.9%[-b]	90.37%[-b]	91.8%[-b]	93.87%[-b]	23.17%[-b]	82.61%[-b]	94.76%
Rare100	63.08%[-b]	60.16%[-b]	60.81%[-b]	64.27%[-b]	30.95%[-b]	39.92%[-b]	52.3%[-b]	51.42%[-b]	58.15%[-b]	12.07%[-b]	34.74%[-b]	72.27%
StaticIm05												
Split3	86.93%[-b]	78.45%[-b]	84.9%[-b]	89.2%[-b]	52.38%[-b]	49.15%[-b]	80.64%[-b]	73.54%[-b]	88.08%[-b]	2.74%[-b]	39.85%[-b]	93.97%
Move3	84.0%[-b]	78.24%[-b]	81.57%[-b]	86.62%[-b]	42.25%[-b]	31.16%[-b]	76.55%[-b]	64.53%[-b]	84.97%[-b]	0.07%[-Ь]	20.98%[-b]	92.36%
Merge3	90.39%[-b]	87.69%[-b]	89.25%[-b]	92.11%[-b]	59.16%[-b]	65.55%[-b]	85.31%[-b]	78.09%[-b]	91.12%[-b]	3.78%[-b]	54.74%[-b]	96.19%
Split7	81.56%[-b]	72.69%[-b]	79.46%[-b]	84.77%[-b]	41.73%[-b]	45.58%[-b]	76.15%[-b]	65.27%[-b]	85.44%[-b]	9.75%[-b]	34.34%[-b]	89.57%
Move7	75.06%[-b]	72.68%[-b]	71.52%[-b]	80.41%[-b]	36.53%[-b]	20.72%[-b]	66.29%[-b]	39.12%[-b]	79.14%[-b]	0.03%[-b]	10.18%[-b]	85.35%
Merge7	87.33%[-b]	86.85%[-b]	85.27%[-b]	90.43%[-b]	51.63%[-b]	65.48%[-b]	82.04%[-b]	69.28%[-b]	88.46%[-b]	14.44%[-b]	52.73%[-b]	93.96%
Borderline20	96.61%[-b]	99.13%[b]	95.93%[-b]	97.67%[-b]	59.51%[-b]	94.88%[-b]	91.99%[-b]	94.27%[-b]	95.46%[-b]	14.95%[-b]	86.71%[-b]	98.57%
Borderline100	95.99%[-b]	98.77%[*]	94.97%[-b]	97.42%[-b]	58.26%[-b]	92.52%[-b]	90.37%[-b]	93.4%[-b]	95.41%[-b]	7.48%[-b]	77.63%[-b]	98.78%
Rare20	92.83%[-b]	92.21%[-b]	92.33%[-b]	93.43%[-b]	46.1%[-b]	73.74%[-b]	88.25%[-b]	90.36%[-b]	93.63%[-b]	11.8%[-b]	73.76%[-b]	94.54%
Rare100	61.79%[-b]	61.89%[-b]	59.97%[-b]	64.29%[-b]	21.15%[-b]	37.6%[-b]	49.57%[-b]	49.17%[-b]	58.63%[-b]	7.37%[-b]	29.16%[-b]	71.94%
					St	aticIm03						
Split3	79.64%[-b]	74.54%[-b]	76.82%[-b]	83.34%[-b]	33.86%[-b]	40.94%[-b]	67.75%[-b]	64.86%[-b]	82.72%[-b]	0.75%[-b]	21.65%[-b]	88.15%
Move3	74.18%[-b]	76.9%[-b]	70.89%[-b]	78.76%[-b]	26.98%[-b]	13.98%[-b]	58.82%[-b]	45.28%[-b]	77.24%[-b]	0.03%[-b]	6.22%[-b]	83.63%
Merge3	83.59%[-b]	85.47%[-b]	82.05%[-b]	86.97%[-b]	45.54%[-b]	53.7%[-b]	71.45%[-b]	66.63%[-b]	84.98%[-b]	1.14%[-b]	30.72%[-b]	91.76%
Split7	72.97%[-b]	71.19%[-b]	69.91%[-b]	76.93%[-b]	25.94%[-b]	38.05%[-b]	60.16%[-b]	51.67%[-b]	78.2%[-b]	3.78%[-b]	19.01%[-b]	81.65%
Move7	62.67%[-b]	71.82%[-s]	57.33%[-b]	69.13%[-b]	24.09%[-b]	9.58%[-b]	44.13%[-b]	31.13%[-b]	69.21%[-b]	0.02%[-b]	3.98%[-b]	73.56%
Merge7	79.72%[-b]	85.68%[-m]	76.69%[-b]	83.89%[-b]	40.91%[-b]	55.71%[-b]	66.38%[-b]	63.1%[-b]	82.59%[-b]	5.64%[-b]	29.95%[-b]	87.89%
Borderline20	93.86%[-b]	98.9%[b]	93.33%[-b]	96.0%[-b]	56.21%[-b]	88.05%[-b]	85.82%[-b]	88.9%[-b]	91.81%[-b]	4.34%[-b]	62.36%[-b]	96.98%
Borderline100	91.62%[-b]	98.54%[b]	90.6%[-b]	95.28%[-b]	47.88%[-b]	76.88%[-b]	75.17%[-b]	78.49%[-b]	87.32%[-b]	2.18%[-b]	48.63%[-b]	97.23%
Rare20	90.76%[-b]	91.57%[-b]	90.13%[-b]	92.13%[-b]	43.55%[-b]	64.12%[-b]	80.52%[-b]	88.84%[-b]	91.43%[-b]	3.52%[-b]	47.06%[-b]	93.24%
Rare100	58.56%[-b]	63.55%[-b]	56.97%[-b]	62.83%[-b]	19.02%[-b]	33.07%[-b]	43.29%[-b]	45.14%[-b]	58.47%[-b]	2.56%[-b]	16.52%[-b]	69.79%

- Based on the average G-Mean, cells are highlighted in lime / light grey when SMOClust performed better than the corresponding approach and cells are highlighted in orange / dark grey cells when SMOClust performed worse than the corresponding approach. The colour saturation scales with the absolute difference in average G-Mean between the SMOClust and the approach of the column and the saturation reaches the maximum when such difference is $\geq 10\%$. - Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against SMOClust. Presence/absence of the sign "-" in the effect size means that the corresponding approach was worse/better than SMOClust.

4 Results with Real-world Data Streams

This section resents the comprehensive results of the predictive performance of approaches on real-world data streams.

- Correspond to Figure 4 in the paper:
 - Figure 7 presents the difference in average G-Mean (based on thirty runs) of the compared approaches against SMOClust on real-world data streams.
 - Table 7 presents average G-Mean (based on thirty runs) of all approaches on real-world data streams and the A12 effect size results of comparing existing approaches against SMOClust.

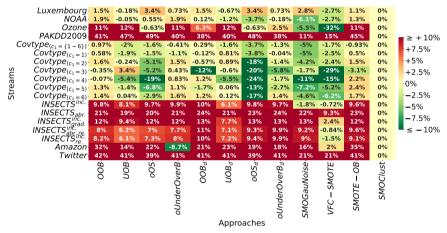


Fig. 7: Difference in Average G-Mean Against SMOClust on Real-World Data Streams Based on 30 Runs (Green cells indicate SMOClust performed better; Red cells indicate SMOClust performed worse)

Stream	OOB	UOB	oOS	oUnder- OverB	$OOB_{\rm d}$	UOB _d	$\mathrm{oOS}_{\mathrm{d}}$	oUnder- OverB _d	SMO- Gau- Noise	VFC- SMOTE	SMOTE- OB	SMO- Clust
Luxembourg	93.29%[b]	91.58%[*]	95.15%[b]	92.49%[s]	93.29%[b]	91.09%[-m]	95.15%[b]	92.49%[s]	94.6%[b]	89.05%[-b]	92.82%[m]	91.76%
NOAA	71.44%[b]	69.47%[s]	70.07%[b]	71.39%[b]	69.64%[*]	68.28%[-b]	65.85%[-b]	69.34%[*]	63.4%[-b]	66.81%[-b]	70.83%[b]	69.52%
Ozone	65.7%[b]	66.49%[b]	54.03%[-m]	65.58%[b]	60.92%[b]	66.44%[b]	54.03%[-m]	57.12%[b]	49.17%[-b]	22.59%[-b]	65.93%[b]	54.66%
PAKDD2009	50.84%[b]	56.8%[b]	57.91%[b]	49.64%[b]	47.56%[b]	49.46%[b]	57.2%[b]	47.8%[b]	20.84%[b]	24.36%[b]	54.21%[b]	9.36%
$Covtype_{(c_1=\{1-6\})}$	91.49%[b]	88.55%[-b]	88.93%[-b]	90.11%[-b]	90.81%[b]	88.91%[-b]	86.84%[-b]	89.19%[-b]	85.47%[-b]	88.83%[-b]	89.59%[-b]	90.52%
$Covtype_{(c_1=1)}$	90.59%[b]	88.1%[-b]	88.51%[-b]	88.91%[-b]	89.89%[-b]	90.82%[b]	86.2%[-b]	89.97%[*]	85.0%[-b]	87.52%[-b]	90.51%[b]	90.01%
$Covtype_{(c_1=2)}$	67.14%[b]	65.32%[-b]	60.45%[-b]	67.08%[b]	64.99%[-b]	66.45%[b]	47.65%[-b]	64.16%[-b]	61.38%[-b]	63.13%[-b]	67.07%[b]	65.56%
$Covtype_{(c_1=3)}$	56.88%[-m]	60.58%[b]	52.02%[-b]	57.66%[b]	45.26%[-b]	56.63%[-b]	37.73%[-b]	51.46%[-b]	55.49%[-b]	28.28%[-b]	54.11%[-b]	57.23%
$Covtype_{(c_1=4)}$	90.05%[-s]	84.73%[-b]	71.16%[-b]	90.95%[b]	91.31%[b]	84.67%[-b]	65.71%[-b]	88.47%[-b]	79.16%[-b]	75.42%[-b]	92.34%[b]	90.12%
$Covtype_{(c_1=5)}$	65.98%[b]	63.22%[-b]	57.8%[-b]	65.74%[b]	62.93%[-b]	64.7%[m]	51.36%[-b]	61.95%[-b]	57.47%[-b]	59.42%[-b]	67.01%[b]	64.64%
$Covtype_{(c_1=6)}$	68.76%[b]	67.44%[s]	64.53%[-b]	69.01%[b]	68.65%[b]	67.52%[b]	50.34%[-b]	68.84%[b]	62.78%[-b]	61.25%[-b]	69.07%[b]	67.4%
INSECTS ^{inc.}	74.91%[b]	73.22%[b]	74.8%[b]	74.95%[b]	75.33%[b]	71.2%[b]	74.87%[b]	74.77%[b]	63.31%[-b]	64.37%[-m]	74.69%[b]	65.09%
INSECTS _{abr.}	73.21%[b]	70.56%[b]	71.65%[b]	72.91%[b]	75.98%[b]	72.91%[b]	74.76%[b]	75.68%[b]	73.54%[b]	61.31%[b]	75.16%[b]	52.0%
INSECTS ^{inc.}	76.96%[b]	74.43%[b]	77.19%[b]	76.78%[b]	77.93%[b]	72.73%[b]	77.53%[b]	78.14%[b]	77.65%[b]	67.42%[b]	76.75%[b]	65.0%
INSECTS ^{inc.} re.	72.11%[b]	70.31%[b]	71.07%[b]	71.81%[b]	74.78%[b]	71.23%[b]	73.35%[b]	74.01%[b]	73.29%[b]	63.25%[-b]	73.68%[b]	64.09%
INSECTS ^{inc.}	72.75%[b]	70.64%[b]	71.78%[b]	72.49%[b]	75.0%[b]	71.67%[b]	73.88%[b]	74.4%[b]	73.49%[b]	62.99%[-m]	73.63%[b]	64.51%
Amazon	63.24%[b]	45.44%[b]	52.82%[b]	22.37%[-b]	52.54%[b]	54.35%[b]	49.59%[b]	48.61%[b]	46.84%[b]	33.07%[s]	66.26%[b]	31.05%
Twitter	63.69%[b]	62.72%[b]	60.79%[b]	63.09%[b]	63.42%[b]	62.94%[b]	60.66%[b]	62.92%[b]	43.08%[b]	42.85%[b]	62.72%[b]	21.95%

Table 7: 30 Runs Average Prequential G-Mean on Real-World Data Streams (A12 SMOClust vs Others)

- Based on the average G-Mean, cells are highlighted in lime / light grey when SMOClust performed better than the corresponding approach and cells are highlighted in orange / dark grey cells when SMOClust performed worse than the corresponding approach. The colour saturation scales with the absolute difference in average G-Mean between the SMOClust and the approach of the column and the saturation reaches the maximum when such difference is > 10%.

Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against SMOClust. Presence/absence of the sign "-" in the effect size means that the corresponding approach was worse/better than SMOClust.