Supplementary material for ECML 2019 submission of paper: An ensemble for classification in multi-class streams with class-based concept drift

Robert Anderson, Yun Sing Koh, and Gillian Dobbie

School of Computer Science, University of Auckland, New Zealand

1 Introduction

This document includes supplementary material to support the submission of 'An ensemble for classification in multi-class streams with class-based concept drift'.

2 Real-world datasets

In our paper, we use popular real-world stream datasets CovType and Poker from `http://moa.cms.waikato.ac.nz/datasets` , Intrusion from `http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html`, HT-Sensor from `http://archive.ics.uci.edu/ml/datasets/gas+sensors+for+home+activity+monitoring` [?], Sensor from `http://www.cse.fau.edu/~xqzhu/stream.html` and include UKCrash [?] a dataset based on severity of one million road accidents in the UK from 2010-2016.

3 Extended results

3.1 CSE performance on real datasets including memory

Table 1 shows Kappa, Temporal-Kappa and memory use for different frameworks on real-world datasets tested.

3.2 Impact of approximating with quantile sketches using severity metric

Table 2 shows the impact of using compact quantile sketches vs. storing all data on *sev*. Changes used in the dataset are as follows: When *before* is less than k, each QS is a perfect representation of the data, but is likely to become more inexact when storing more samples. As we use k rather than samples seen to calculate p-values in *sev*, using QS's should always result in more conservative *sev*, which is supported by these results. However, there is a strong relationship between *sev* when using and not using QSs across all datasets.

Table 1: Kappa, Temporal-Kappa and memory for ensemble frameworks on real-world datasets

Stream	Kappa					Kappa-Temporal					Memory (000,000s of bytes)				
	CSE	AUE	DP	Dynse	OB Adwin	CSE	AUE	DP	Dynse	OB Adwin	CSE	AUE	DP	Dynse	OB Adwin
Covtype	0.41	0.55	0.36	**0.72**	0.45	-6.6	-4.5	-6.8	**-2.4**	-5.6	1.4	**0.4**	1.7	3.4	0.3
HT˙Sensor	**0.23**	0.06	0.01	0.18	0.06	**0.2**	0.1	0.0	**0.2**	0.1	9.0	**0.3**	0.5	2.0	2.0
Intrusion	0.93	0.00	0.96	**0.98**	0.96	-35.1	-741.1	-20.2	**-11.6**	-22.8	6.4	**1.4**	1.5	2.7	1.0
Poker	0.30	0.27	0.33	**0.43**	0.35	-0.5	-0.5	-0.4	**-0.2**	-0.3	0.6	**0.2**	0.5	1.4	**0.2**
Sensor	0.13	0.22	0.07	**0.59**	0.18	0.1	0.2	0.1	**0.6**	0.2	2.9	0.6	1.4	1.0	**0.2**
UKCrash	**0.01**	0.00	0.00	0.00	0.00	-0.5	**0.4**	**0.4**	**0.4**	**0.4**	**1.5**	27.4	12.0	1.9	27.8

Table 2: Impact of Quantile Sketch approximation (QS) on *sev* in presence of change Δ and no change with varied *before* size ($k = 256$, $after$ size of 128 instances)

before size	Circles (Δ: 0.2→0.4)				MixedType (Δ: 10→6)				RBF (Δ: 50→150)				RUnif (Δ: 0.2→0.8)			
	No QS No Δ	No QS Δ	QS No Δ	QS Δ	No QS No Δ	No QS Δ	QS No Δ	QS Δ	No QS No Δ	No QS Δ	QS No Δ	QS Δ	No QS No Δ	No QS Δ	QS No Δ	QS Δ
100	0.766	0.091	0.766	0.091	0.989	0.036	0.989	0.036	0.826	0.419	0.826	0.419	0.800	0.212	0.800	0.212
500	0.722	0.010	0.722	0.010	0.931	0.004	0.931	0.004	0.843	0.188	0.847	0.197	0.728	0.035	0.729	0.036
1000	0.778	0.005	0.833	0.009	0.899	0.005	0.899	0.005	0.873	0.110	0.965	0.195	0.743	0.030	0.824	0.050
5000	0.801	0.003	0.891	0.010	0.924	0.006	0.939	0.011	0.849	0.090	0.969	0.266	0.753	0.012	0.862	0.040
10000	0.764	0.003	0.892	0.010	0.931	0.005	0.955	0.010	0.849	0.087	0.970	0.276	0.765	0.014	0.877	0.044

Table 3: *sev* with perfectly correlated attributes in $before$ and/or $after$ windows when using KS vs. CVM ($before/after$ size of 64/64)

Change	Test	Correlated attributes: Neither	*before*	*after*	Both
No Δ	CVM	0.506 (0.285)	0.000 (0.000)	0.000 (0.000)	0.516 (0.293)
No Δ	KS	0.512 (0.306)	0.578 (0.308)	0.649 (0.327)	0.524 (0.286)
Δ	CVM	0.074 (0.140)	0.000 (0.000)	0.000 (0.000)	0.370 (0.251)
Δ	KS	0.216 (0.209)	0.389 (0.325)	0.438 (0.276)	0.469 (0.288)

3.3 Impact of correlated attributes

In Table 3, we compare our severity metric using aggregated KS comparisons with using CVM on 64 instance $before$ and $after$ windows of instances with ten uniformly randomly generated attributes $\in [0, 1]$. We show our measure with no change and a change of $+0.02$ in the *after* window. When attributes are correlated $before$ and $after$ window impact *sev* to examine what information we lose through not considering conditional probabilities in univariate tests. CVM uses conditional probabilities to identify where a the windows are different while with univariate tests, our measure does not. KS is notably more conservative than CVM in cases with change.

References