Early and Revocable Time Series Classification: Supplementary material

This document gives further details on experiments presented in our paper, Section 1 shows that the percentage of samples for which a revocation is useful is relatively low, that demonstrates that the tackled problem is hard, then Section 2 presents results of the Friedmann test comparing the two proposed approaches, Section 3 shows the pareto curves for additional values of β and finally experiments on multi-class classification problems are presented in Section 4.

1 Percentage of samples for which a revocation is useful

This section is an extension to Section 5.5 in our paper, In order to check if the cost paid by the user can be improved by revoking the decision made by the irrevocable regime, statistics about results of ECONOMY-γ were computed. Figure 6 shows the percentage of samples for which the score paid by the user can be improved by revoking the decision of ECONOMY-γ for each dataset and α value.

These statistics show that decisions can be successfully revoked on very few samples which makes the *early and revocable time series classification* problem difficult to solve, because the algorithm needs to identify those samples using incomplete knowledge of the time series in order to revoke the decision on them and keep the decision of the irrevocable regime on the other samples.

Datasets	0.0001	0.0003	0.0005	0.0008	0.001	0.0025	0.005	0.0075	0.01	0.025	0.05	0.075	0.1	0.25	0.5	0.75	1
CBF	1.075	1.075	1.075	1.075	1.075	1.075	1.075	1.075	1.075	1.075	1.075	2.151	3.226	3.226	4.659	5.018	6.452
ChlorineConcentration	12.142	12.142	12.142	12.142	12.142	12.142	12.142	12.142	12.142	16.705	12.142	18.097	20.65	21.036	21.346	23.279	23.511
CinCECGTorso	1.174	1.174	1.174	1.174	1.174	1.174	1.174	1.174	1.174	1.174	1.174	1.174	1.174	0.939	1.643	1.643	1.643
Crop	0.111	0.111	0.111	0.111	0.111	0.139	0.139	0.139	0.139	0.153	0.417	0.486	0.5	1.014	3.681	3.681	3.681
ECG5000	0.267	0.267	0.267	0.267	0.267	0.267	0.533	0.6	0.8	0.867	0.867	0.867	0.933	0.933	1.667	1.6	1.733
ECGFiveDays	0	0	0	0	0	0	0	0	0	0	2.256	2.256	2.256	2.256	10.902	8.271	12.406
ElectricDevices	0	0	0	0	0	0	0	0	0	1.302	1.302	5.028	5.329	8.594	10.877	12.28	13.241
FaceAll	0	0	0	0	0	0	0	0	0	0	0	0	0.296	0.148	0.296	0.444	0.444
FacesUCR	1.63	1.63	1.63	1.63	1.63	1.63	1.63	1.63	1.63	2.519	2.074	3.407	3.407	4.741	5.778	6.519	6.519
FiftyWords	4.044	4.044	4.044	4.044	4.044	4.044	4.044	4.044	4.044	4.044	2.206	4.779	4.779	3.309	5.515	4.412	5.515
FordA	0	0	0	0	0	0	2.031	2.031	2.031	2.573	2.979	3.453	3.859	4.875	5.281	7.177	10.494
FreezerRegularTrain	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.556	0.556	1.333	1.333
HandOutlines	4.38	4.38	4.38	4.38	4.38	4.623	4.623	4.623	4.623	4.623	6.083	4.623	8.029	8.273	11.922	14.599	15.815
InsectWingbeatSound	1.97	1.97	1.97	1.97	1.97	1.97	4.848	4.848	4.848	4.848	1.667	1.667	2.424	3.333	4.697	5	7.727
ItalyPowerDemand	0.912	0.912	0.912	0.912	0.912	0.912	0.912	0.912	0.912	3.04	3.04	2.432	5.775	8.815	18.541	22.188	25.836
Mallat	0.139	0.139	0.139	0.139	0.139	0.139	0.139	0.139	0.139	0.417	0.417	0.556	0.556	1.528	2.222	3.889	3.611
MedicalImages	9.621	9.621	9.621	9.621	9.621	9.621	9.621	9.621	9.621	9.913	9.913	11.662	12.245	18.659	18.659	23.907	23.907
MelbournePedestrian	0.367	0.367	0.367	0.367	0.367	0.367	0.367	0.367	0.367	0.642	0.642	0.826	0.826	1.743	1.651	2.294	8.532
MixedShapesRegularTrain	1.025	1.025	1.025	1.025	1.025	1.025	1.025	1.025	1.025	1.253	1.253	1.253	1.367	1.822	1.822	2.733	2.164
MoteStrain	3.141	3.141	3.141	3.141	3.141	3.141	3.141	2.094	2.094	3.927	3.927	5.236	6.021	6.806	7.068	8.377	15.445
NonInvasiveFatalECGThorax2	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.708	1.239
PhalangesOutlinesCorrect	15.414	15.414	15.414	15.414	15.414	15.414	15.414	16.165	16.165	14.787	18.045	18.045	18.045	22.306	24.311	28.947	30.827
ProximalPhalanxOutlineCorrect	8.209	8.209	8.209	8.209	8.209	8.209	8.209	8.209	8.209	9.701	9.701	10.821	11.194	17.164	26.493	26.119	26.493
SemgHandGenderCh2	7	7	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	7	11	15	14.5
SonyAIBORobotSurface2	5.442	5.442	5.442	5.442	5.442	5.442	5.442	5.442	5.442	5.442	5.442	5.442	5.442	6.463	8.503	10.204	9.864
StarLightCurves	0	0	0	0	0	0	0.974	1.155	1.191	1.119	1.263	2.382	2.382	2.382	3.428	4.15	4.367
Strawberry	4.407	4.407	4.407	4.407	4.407	4.407	4.407	4.407	4.746	4.746	5.085	5.085	5.085	7.119	13.559	14.237	15.593
Symbols	1.333	1.333	1.333	1.333	1.333	1.333	1.333	1.333	1.333	1.333	1.333	2	2.333	2.333	3	3	10
TwoLeadECG	0.86	0.86	0.86	0.86	0.287	0.287	0.287	0.287	0.86	0.86	0.86	0.86	0.86	2.292	4.585	5.158	8.023
TwoPatterns	1.467	1.467	1.467	1.467	1.467	1.467	1.467	1.467	1.867	1.933	1.933	3.133	3.267	10.067	24.533	25.067	25.4
UWaveGestureLibraryX	1.265	1.265	1.265	1.265	1.265	1.265	1.265	1.265	1.265	1.562	2.455	2.455	2.679	3.646	4.836	6.994	7.887
Wafer	0	0	0	0	0	0.047	0.047	0.047	0.047	0.047	0.14	0.14	0.14	0.14	0.186	0.279	0.279
WordSynonyms	5.515	5.515	5.515	5.515	5.515	5.515	5.515	5.515	5.515	5.882	4.412	6.985	4.412	6.25	8.456	9.559	14.338
Yoga	10.303	10.303	10.303	10.303	10.303	10.303	10.505	10.808	10.808	10.808	11.616	12.626	12.424	13.232	13.434	14.747	14.747

Figure 1: Percentage of samples for which the score paid by the user can be improved by revoking the decision of ECONOMY-γ. Each column represent a different value of α in ascending order.

2 Average ranking using the Friedman test

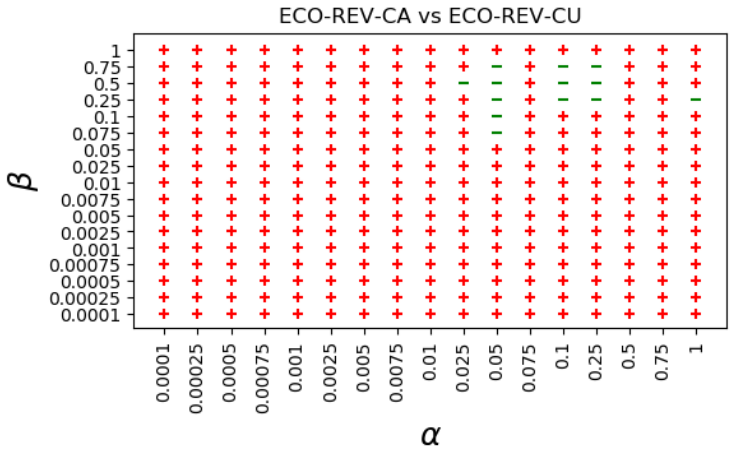

Figure 2: ECO-REV-CA *vs.* ECO-REV-CU: Average ranking based on *AvgCost* using the Friedman test, for different values of α and β: "+" indicates ECO-REV-CA having a better average rank than ECO-REV-CA, and "−" indicates the opposite .

In our paper, we used Wilcoxon signed-rank test to compare the two proposed approaches (see Section 5). In addition, Figure 2 uses the Friedman test and shows that ECO-REV-CA has better average rank than ECO-REV-CU in 96% of cases.

3 Average Earliness vs. Average Kappa for different values of β

This section presents the pareto curve for additional values of β, we notice that the same order holds for all values of β except for β=0.75 and β=1 where ECO-REV-CA becomes worse than ECONOMY-γ, however ECO-REV-CU still performs better.

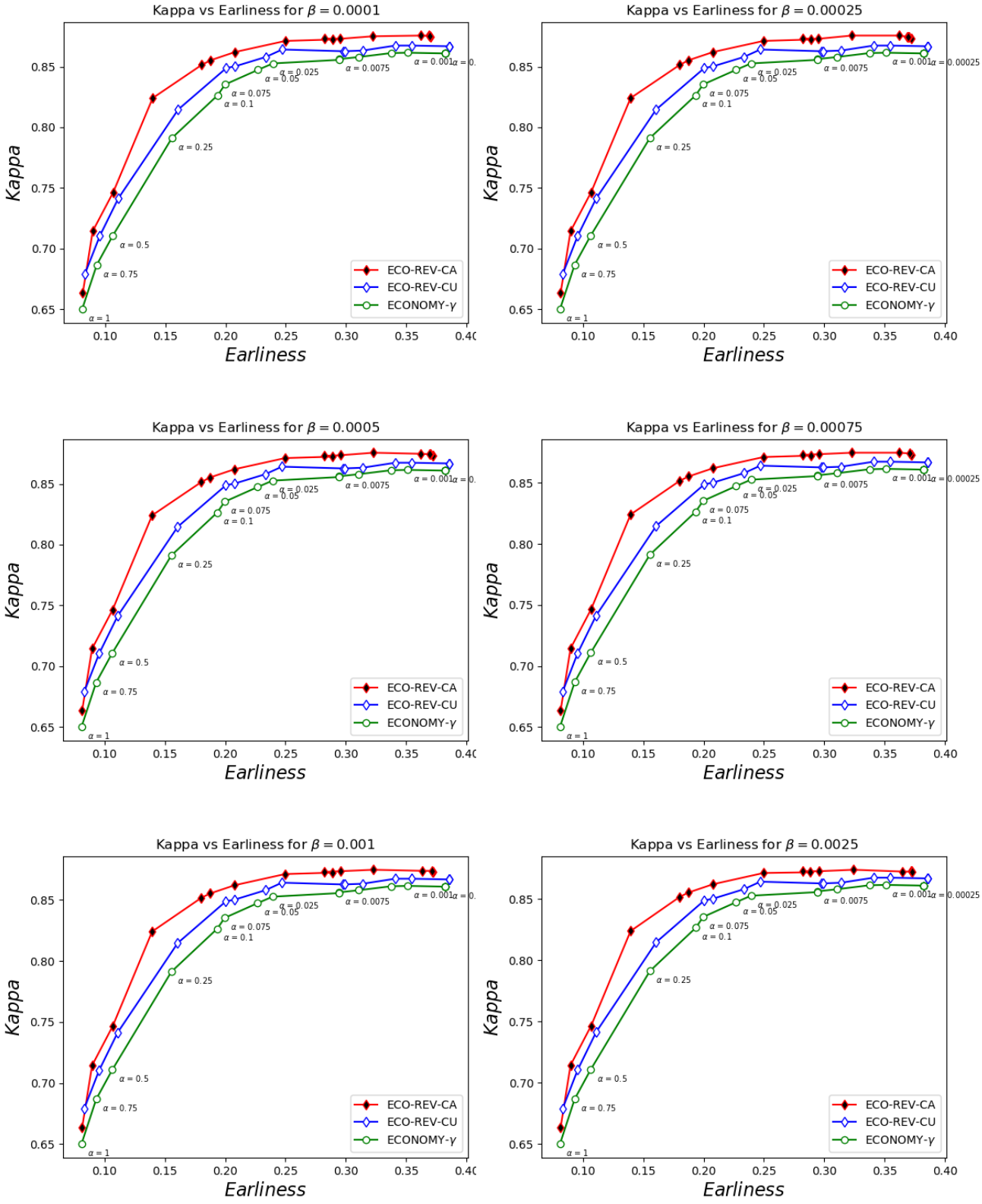

Figure 3: Average *Earliness* vs. Average *Kappa* score obtained over the 34 datasets by varying α of the delay cost.

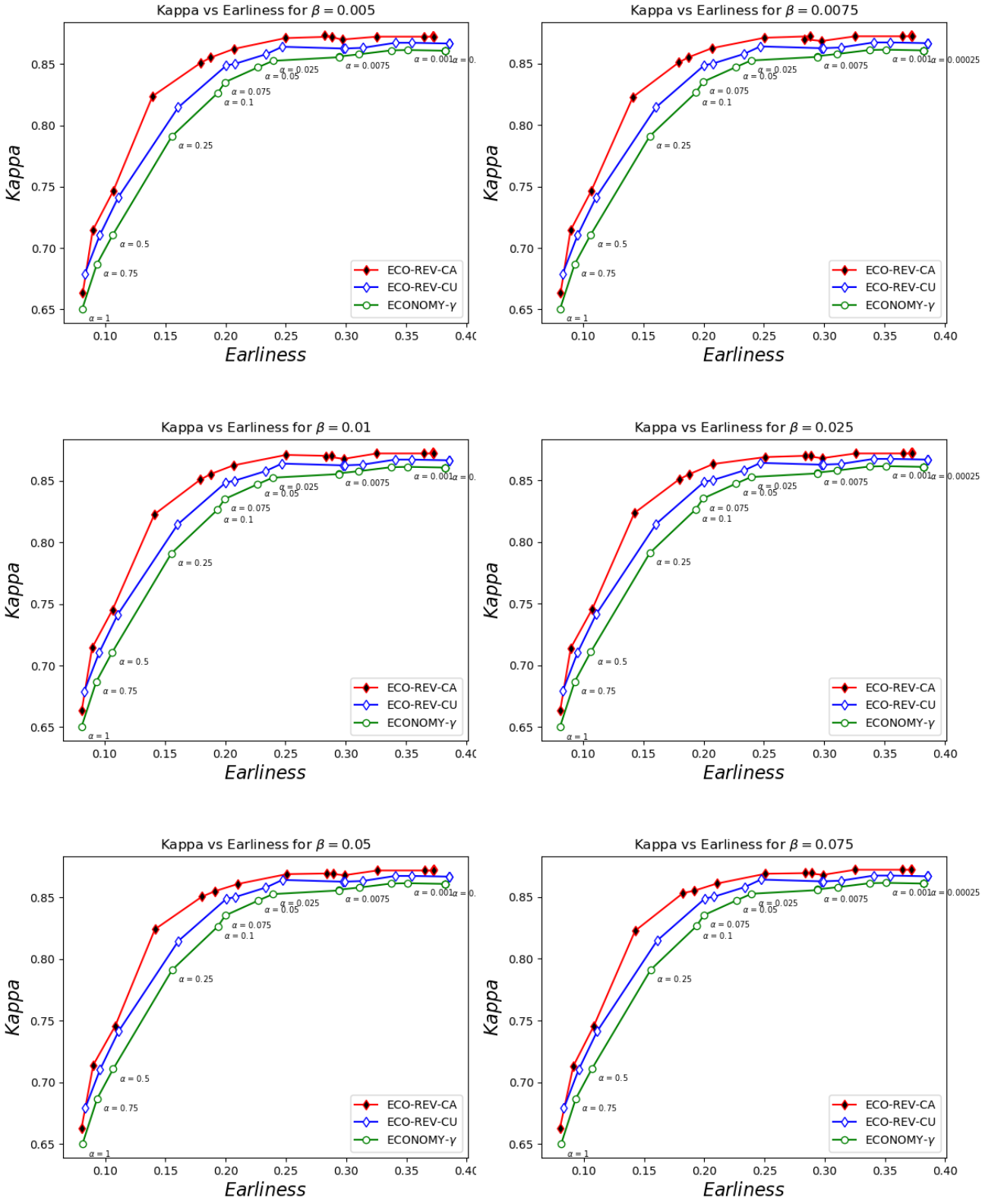

Figure 4: Average *Earliness* vs. Average *Kappa* score obtained over the 34 datasets by varying α of the delay cost.

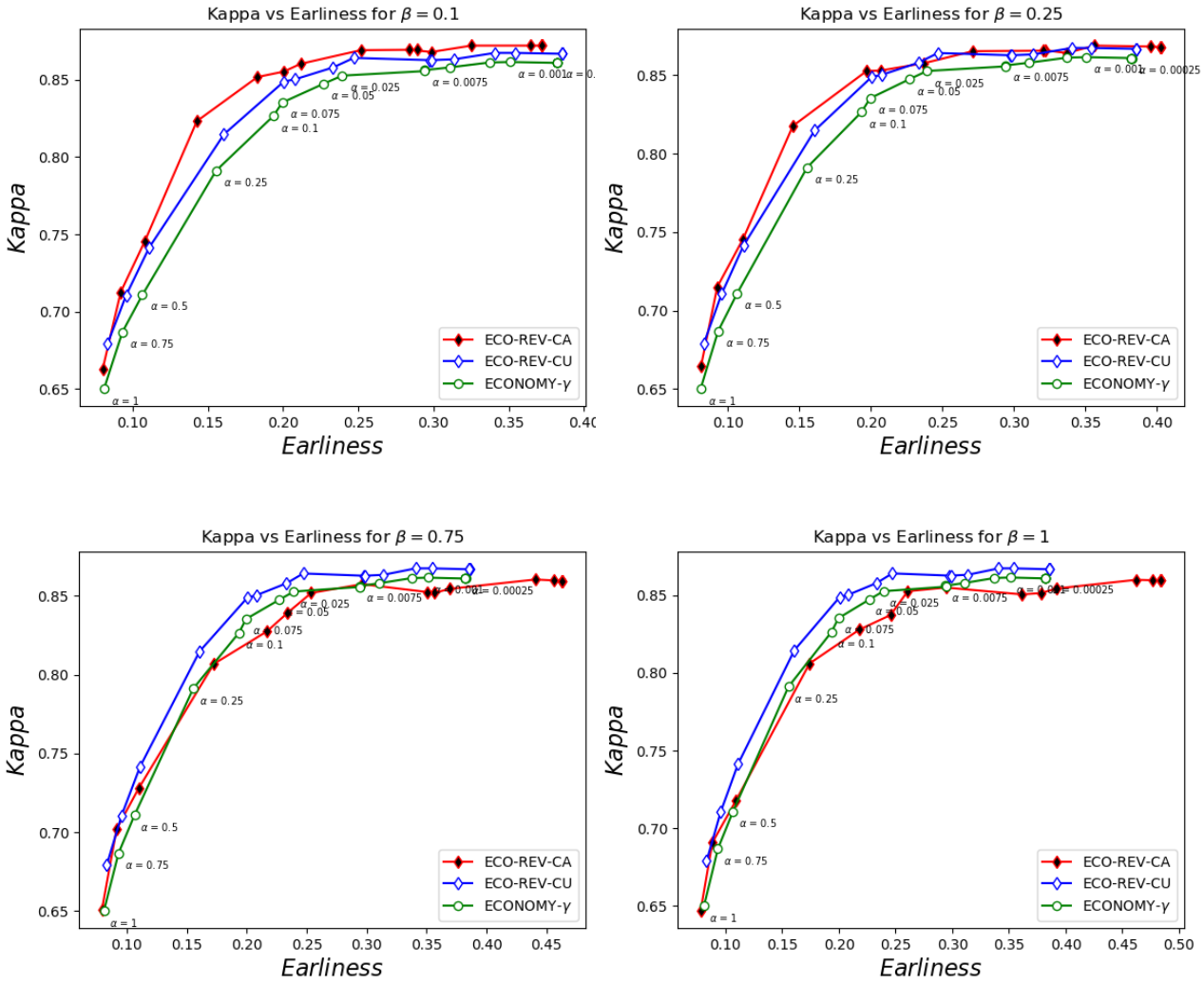

Figure 5: Average *Earliness* vs. Average *Kappa* score obtained over the 34 datasets by varying α of the delay cost.

4 Experiments on multi-class early and revocable classification problems

This section gives two simple implementation examples of the ECONOMY framework, in the case of multi-class problems and revocable decisions.

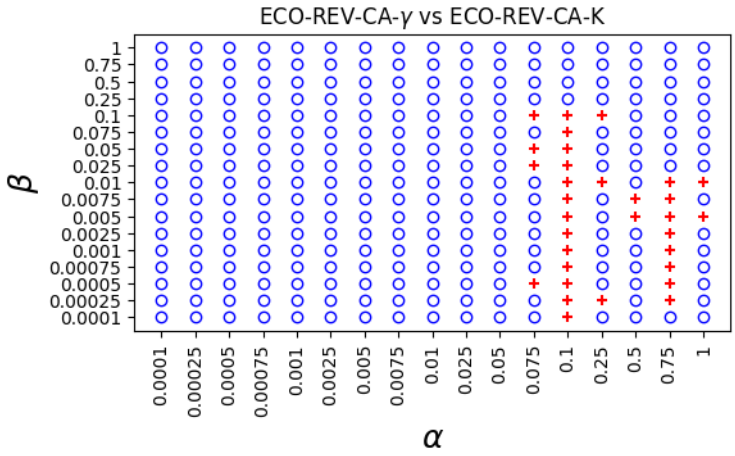

Figure 6: ECO-REV-CA-γ vs. ECO-REV-CA-K: Wilcoxon signed-rank test applied on the *AvgCost* criterion over the 22 test sets, for a range of couples of values α and β, with "+" indicating a significant success of the first approach, "∘" an insignificant difference and "−" indicating a significant failure of the first approach.

ECONOMY-γ is by design limited to binary early classification problems. The reason behind this limitation is that the partitioning of time series in the training set is done by discretising the confidence level of the classifiers regarding the positive class. We propose a new method ECONOMY-γ-MC which handles multi-class early classification problems. The intuition behind this method is that time series with similar uncertainty level of classifiers belong to the same group. In fact, instead of discretising the confidence level of the classifiers to the positive class, we propose to discretise the uncertainty of the classifier. This uncertainty is measured using the relative entropy function. The mechanism of predicting the groups at future time steps using Markov chain remains unchanged compared to ECONOMY-γ.

ECONOMY-K partition time series using the clustering algorithm K-means. This approach can deal by design with multi-class classification problems. ECONOMY-K and ECONOMY-γ-MC are used to implement early and revocable framework described in our paper. Experiments were performed on 22 datasets chosen at random from the UAE-UCR archive to compare performance of these two revocable strategies called respectively ECO-REV-CA-K and ECO-REV-CA-γ. Results of Wilcoxon signed-rank test applied on the *AvgCost* criterion are shown in Figure 6. Further detailed results focusing on *AvgCost* are presented in Table 1 and Table 2.

Datasets	**Class**	α **= 0.0001**		α **= 0.0025**		α **= 0.0005**		α **= 0.00075**		α **= 0.001**		α **= 0.0025**		α **= 0.005**		α **= 0.0075**		α **= 0.01**	
		G	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**
Plane	7	0.0953	0.0955	0.0953	0.0956	0.0953	0.0957	0.0953	0.0959	0.0954	0.096	0.0956	0.0969	0.0959	0.0827	0.0962	0.0841	0.0965	0.0694
BirdChicken	2	0.3334	0.3334	0.3335	0.3335	0.3337	0.3337	0.3338	0.3338	0.334	0.334	0.335	0.335	0.3368	0.3368	0.3385	0.3385	0.3402	0.3402
MoteStrain	2	0.0341	0.0446	0.0342	0.0473	0.0343	0.0475	0.0344	0.0477	0.0344	0.0479	0.0349	0.0361	0.0357	0.0381	0.0366	0.0401	0.0374	0.0395
ProximalPhalanxOutlineCorrect	2	0.1197	0.1158	0.1198	0.1159	0.1199	0.1162	0.1201	0.1164	0.1239	0.1166	0.1246	0.1181	0.1258	0.1205	0.127	0.1228	0.1352	0.1252
Fungi	18	0.1775	0.0807	0.1775	0.0808	0.1776	0.081	0.1777	0.0811	0.1778	0.0813	0.1785	0.0822	0.1795	0.0838	0.1806	0.117	0.1816	0.1184
RefrigerationDevices	3	0.4134	0.4134	0.4134	0.4134	0.4136	0.4136	0.4137	0.4137	0.4138	0.4138	0.4145	0.4145	0.4156	0.4156	0.4167	0.4167	0.3597	0.4178
SemgHandMovementCh2	6	0.2993	0.3038	0.2994	0.3039	0.2995	0.3041	0.2996	0.3043	0.2998	0.3046	0.3005	0.3058	0.3017	0.308	0.3069	0.3101	0.3081	0.3122
Ham	2	0.2785	0.4154	0.2785	0.4155	0.2786	0.4156	0.2786	0.4157	0.2787	0.4158	0.279	0.4164	0.2799	0.4173	0.2804	0.4183	0.2809	0.4193
Wine	2	0.3824	0.1765	0.3824	0.1766	0.3825	0.1768	0.3826	0.1769	0.3827	0.1771	0.3831	0.178	0.3838	0.1795	0.3846	0.1811	0.3853	0.1825
HandOutlines	2	0.1193	0.0949	0.1193	0.095	0.1194	0.0951	0.1195	0.0952	0.1196	0.0953	0.1201	0.0959	0.1209	0.0969	0.1216	0.0979	0.1003	0.0989
OliveOil	4	0.2778	0.5001	0.2778	0.5002	0.2779	0.5004	0.2779	0.5007	0.278	0.5009	0.2782	0.5022	0.2787	0.5045	0.2792	0.5067	0.2797	0.5087
FreezerRegularTrain	2	0.0011	0.0023	0.0011	0.0023	0.0012	0.0024	0.0012	0.0024	0.0012	0.0014	0.0014	0.0028	0.0017	0.0033	0.002	0.0047	0.0022	0.0061
Earthquakes	2	0.2015	0.2159	0.2015	0.216	0.2016	0.2161	0.2016	0.2163	0.2017	0.2164	0.2021	0.2173	0.2027	0.2188	0.2031	0.2202	0.2036	0.2217
Trace	4	0.0334	0.0334	0.0335	0.0336	0.0336	0.0338	0.0337	0.034	0.0338	0.0342	0.0345	0.0356	0.0357	0.0378	0.0369	0.0401	0.0381	0.0423
SonyAIBORobotSurface1	2	0.0321	0.0268	0.0322	0.0269	0.0324	0.0271	0.0326	0.0273	0.0327	0.0275	0.0337	0.0233	0.0353	0.0301	0.0369	0.027	0.0385	0.033
ArrowHead	3	0.2188	0.1719	0.2188	0.1721	0.2189	0.1722	0.219	0.1724	0.219	0.1726	0.2195	0.1737	0.2202	0.1756	0.2209	0.1775	0.2216	0.1793
ChlorineConcentration	3	0.219	0.2174	0.2191	0.2175	0.2193	0.2178	0.2195	0.2172	0.2197	0.2174	0.221	0.2187	0.2181	0.2209	0.219	0.217	0.2261	0.2252
Strawberry	2	0.0611	0.0511	0.0611	0.0512	0.0612	0.0513	0.0613	0.0515	0.0614	0.0516	0.0619	0.0557	0.0659	0.064	0.0666	0.0783	0.0673	0.0784
ScreenType	3	0.4408	0.4535	0.4409	0.4537	0.4411	0.4539	0.4413	0.4542	0.4415	0.4544	0.4428	0.4514	0.4448	0.4538	0.4469	0.4695	0.449	0.4761
InsectEPGRegularTrain	3	0.0	0.0	0.0	0.0001	0.0	0.0001	0.0	0.0002	0.0	0.0003	0.0001	0.0006	0.0002	0.0011	0.0004	0.012	0.0005	0.0123
FordA	2	0.3016	0.3016	0.3016	0.3016	0.3016	0.3016	0.3016	0.3016	0.3016	0.3016	0.3017	0.3017	0.3018	0.3018	0.302	0.302	0.3021	0.3021
FacesUCR	14	0.0617	0.1215	0.0617	0.1216	0.0618	0.1218	0.0619	0.122	0.062	0.1221	0.0626	0.1231	0.0635	0.1247	0.0601	0.1263	0.061	0.1279

Table 1: AvgCost for every dataset in the 22 benchmark for ECO-REV-CA-γ and ECO-REV-CA-K given a fixed value of $\beta = 0.01$

Datasets	**Class**	$\alpha = 0.025$		$\alpha = 0.05$		$\alpha = 0.075$		$\alpha = 0.1$		$\alpha = 0.25$		$\alpha = 0.5$		$\alpha = 0.75$		$\alpha = 1$	
		G	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**	**G**	**K**
Plane	7	0.0985	0.074	0.1017	0.0664	0.1049	0.0736	0.1081	0.0641	0.1275	0.1427	0.1597	0.1792	0.2093	0.2698	0.2124	0.2764
BirdChicken	2	0.3504	0.3504	0.3675	0.3678	0.3846	0.3833	0.4017	0.4822	0.5042	0.5042	0.6851	0.6322	0.7359	0.6984	0.6709	0.6533
MoteStrain	2	0.0518	0.0565	0.0606	0.0821	0.0595	0.097	0.0772	0.1061	0.1236	0.1598	0.169	0.1917	0.1915	0.2124	0.2162	0.2248
ProximalPhalanxOutlineCorrect	2	0.1415	0.1394	0.1249	0.1868	0.1524	0.2204	0.1686	0.2294	0.1994	0.2697	0.2601	0.316	0.3024	0.3285	0.3239	0.341
Fungi	18	0.1879	0.1266	0.1984	0.1403	0.2088	0.1539	0.2193	0.1676	0.2821	0.2497	0.4098	0.3758	0.4782	0.5281	0.6025	0.6529
RefrigerationDevices	3	0.3627	0.4246	0.3727	0.4358	0.3724	0.4471	0.3886	0.4583	0.4075	0.4331	0.4236	0.4706	0.4283	0.4691	0.4533	0.4848
SemgHandMovementCh2	6	0.3149	0.325	0.3263	0.3462	0.3253	0.3943	0.3491	0.3741	0.3922	0.4854	0.4827	0.565	0.6231	0.6231	0.6481	0.6481
Ham	2	0.2842	0.4251	0.2896	0.4349	0.295	0.4446	0.3004	0.4544	0.2993	0.5388	0.3803	0.4771	0.4249	0.5156	0.4518	0.5629
Wine	2	0.3898	0.1914	0.3972	0.1764	0.4046	0.191	0.4121	0.2057	0.4567	0.3168	0.531	0.4769	0.4531	0.669	0.5844	0.6403
HandOutlines	2	0.1045	0.1049	0.1114	0.1148	0.1153	0.1248	0.122	0.1348	0.1655	0.1754	0.2052	0.2359	0.2366	0.2791	0.2595	0.2853
OliveOil	4	0.2825	0.4085	0.2872	0.3661	0.2919	0.3236	0.2966	0.3226	0.3249	0.3218	0.3719	0.3569	0.419	0.3842	0.4661	0.3799
FreezerRegularTrain	2	0.007	0.0064	0.0062	0.0095	0.0074	0.0122	0.0109	0.0153	0.0212	0.0262	0.0361	0.0383	0.0517	0.0507	0.06	0.0632
Earthquakes	2	0.2087	0.2641	0.2352	0.247	0.2412	0.2483	0.2472	0.2495	0.2568	0.2568	0.269	0.269	0.2812	0.2812	0.2934	0.2934
Trace	4	0.0452	0.0555	0.0571	0.0582	0.069	0.0869	0.0853	0.1123	0.2147	0.1399	0.3681	0.3112	0.3768	0.3162	0.3483	0.3541
SonyAIBORobotSurface1	2	0.0659	0.04	0.0845	0.0643	0.065	0.0503	0.067	0.075	0.0801	0.1131	0.1016	0.1284	0.1085	0.1391	0.1412	0.1498
ArrowHead	3	0.226	0.1903	0.2332	0.1762	0.2405	0.2082	0.1584	0.2054	0.2912	0.2949	0.3071	0.301	0.2977	0.4181	0.4369	0.4618
ChlorineConcentration	3	0.2335	0.2404	0.2388	0.2551	0.2474	0.2743	0.2545	0.2871	0.2923	0.3285	0.3425	0.3943	0.384	0.4239	0.4077	0.4456
Strawberry	2	0.0749	0.0754	0.0824	0.0859	0.0844	0.0895	0.09	0.1022	0.1254	0.1504	0.1718	0.2129	0.213	0.2396	0.2371	0.2846
ScreenType	3	0.4635	0.4524	0.477	0.4813	0.4718	0.4817	0.489	0.5051	0.5684	0.6107	0.6575	0.6345	0.6826	0.6429	0.6879	0.6462
InsectEPGRegularTrain	3	0.0012	0.0151	0.0025	0.0059	0.0037	0.0085	0.005	0.0112	0.0125	0.0125	0.025	0.025	0.0374	0.0374	0.0499	0.0499
FordA	2	0.3028	0.3028	0.3041	0.3041	0.3053	0.3053	0.3066	0.3066	0.3141	0.3141	0.3266	0.3266	0.3391	0.3391	0.3516	0.3516
FacesUCR	14	0.0665	0.1375	0.0829	0.1535	0.0968	0.1696	0.1031	0.1856	0.1638	0.2826	0.2626	0.3565	0.3352	0.4482	0.3966	0.5166

Table 2: AvgCost for every dataset in the 22 benchmark for ECO-REV-CA-γ and ECO-REV-CA-K given a fixed value of $\beta = 0.01$