

A Supplementary Material: Additional Figures

In this section, we report figures on the cumulative losses for Experiments #1 (Subsection 6.5) and #2 (Subsection 6.6) of the single FDSs over time and on the evolution of the weights of the MWU method.

At first, in Figure 4, we report a detailed legend to list the different ML methods we employed as base FDS. The same colors have been used to distinguish the different base models in the following figures (both about the loss and the MWU weights).

Figures 5 and 6 provide the cumulated loss of all the analysed methods (ensemble, online, and base models) for the setting of Experiments 1 and 2, respectively. Finally, Figures 7 and 8 provide the evolution of the weights of the online learning model over time for Experiments 1 and 2, respectively. This shows how the online approach is required here since the best-performing model changes over time and the weights follow such a behavior.

— LR(LT,IS)	— RF(LT,IS)	— XGB(LT,IS)
— LR(LT,TH)	— RF(LT,TH)	— XGB(LT,TH)
— LR(ST,IS)	— RF(ST,IS)	— XGB(ST,IS)
— LR(ST,TH)	— RF(ST,TH)	— XGB(ST,TH)
— LR(MT,IS)	— RF(MT,IS)	— XGB(MT,IS)
— LR(MT,TH)	— RF(MT,TH)	— XGB(MT,TH)
— NN(LT,IS)	— SVM(LT,IS)	★ MWU
— NN(LT,TH)	— SVM(LT,TH)	★ Mean
— NN(ST,IS)	— SVM(ST,IS)	★ Majority Voting
— NN(ST,TH)	— SVM(ST,TH)	
— NN(MT,IS)	— SVM(MT,IS)	
— NN(MT,TH)	— SVM(MT,TH)	

Fig. 4: Legend for the different ML methods used in the experimental section. We also report the ensemble methods (MEAN and Majority) and our online learning method (MWU) for completeness.

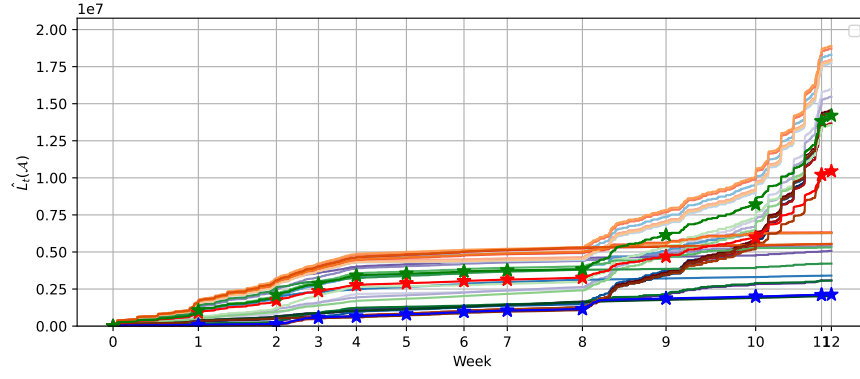


Fig. 5: Cumulative loss for the base models, the ensemble approaches, and the MWU on for Exp. 1. Legend in Figure 4.

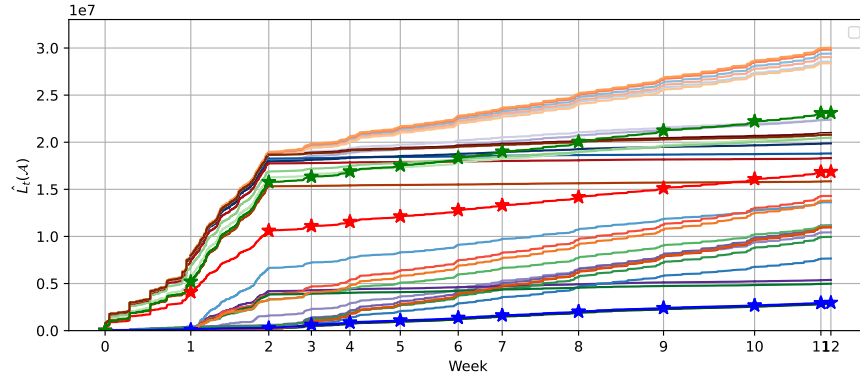


Fig. 6: Cumulative loss for the base models, the ensemble approaches, and the MWU on for Exp. 2. Legend in Figure 4.

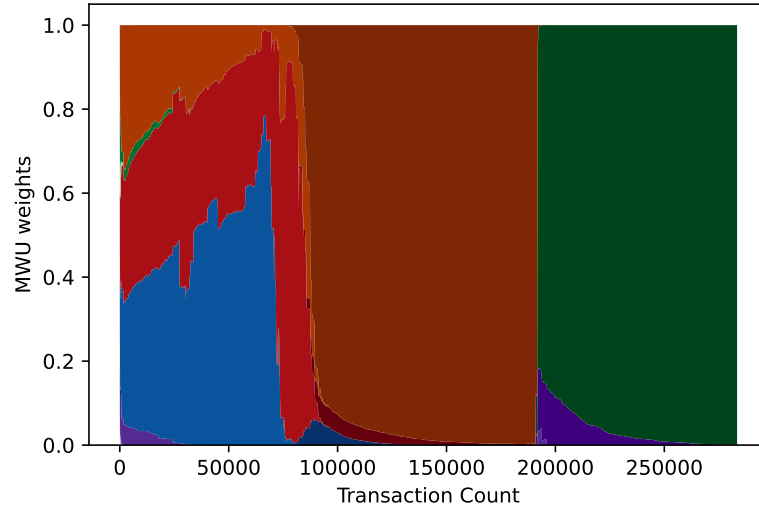


Fig. 7: MWU weights over time for Exp. 1. Legend in Figure 4.

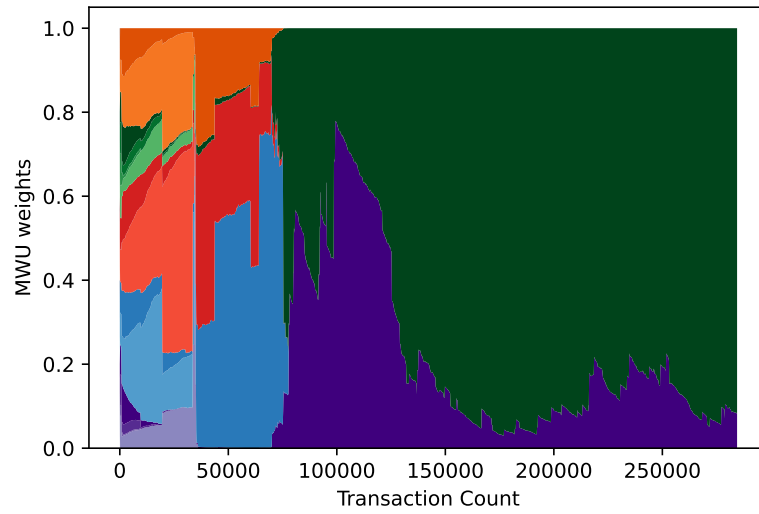


Fig. 8: MWU weights over time for Exp. 2. Legend in Figure 4.

B Base model performances

In this section, we provide Table 4, which contains the full-scale evaluation with the F1-score metric of the performances all of the base models \mathcal{M}_k on each test dataset $\mathcal{D}(r, s)$.

Table 4: F1-Score of model \mathcal{M}_k on each test dataset $\mathcal{D}(r, s)$.

		F1-Score on dataset $\mathcal{D}(r, s)$					
		ST,IS	ST,TH	MT,IS	MT,TH	LT,IS	LT,TH
Logistic Regression	ST,IS	0.70	0.46	0.13	0.03	0.00	0.00
	ST,TH	0.49	0.40	0.00	0.14	0.00	0.00
	MT,IS	0.43	0.21	0.44	0.16	0.28	0.09
	MT,TH	0.26	0.19	0.25	0.14	0.19	0.09
	LT,IS	0.00	0.02	0.33	0.08	0.38	0.17
	LT,TH	0.01	0.03	0.18	0.08	0.20	0.12
Neural Network	ST,IS	0.77	0.32	0.33	0.06	0.00	0.00
	ST,TH	0.43	0.33	0.15	0.21	0.00	0.00
	MT,IS	0.05	0.10	0.64	0.24	0.00	0.00
	MT,TH	0.00	0.10	0.40	0.26	0.00	0.03
	LT,IS	0.00	0.02	0.07	0.01	0.80	0.28
	LT,TH	0.00	0.00	0.00	0.00	0.28	0.35
Random Forest	ST,IS	0.73	0.39	0.00	0.00	0.00	0.00
	ST,TH	0.43	0.33	0.00	0.00	0.00	0.00
	MT,IS	0.06	0.06	0.61	0.28	0.00	0.00
	MT,TH	0.00	0.00	0.49	0.37	0.00	0.04
	LT,IS	0.00	0.00	0.01	0.00	0.67	0.27
	LT,TH	0.00	0.00	0.00	0.00	0.14	0.24
Support Vector Machine	ST,IS	0.78	0.48	0.26	0.05	0.00	0.00
	ST,TH	0.53	0.44	0.06	0.23	0.00	0.00
	MT,IS	0.46	0.22	0.47	0.14	0.33	0.09
	MT,TH	0.30	0.22	0.28	0.16	0.17	0.07
	LT,IS	0.15	0.01	0.34	0.03	0.45	0.08
	LT,TH	0.05	0.04	0.18	0.07	0.20	0.11
Extreme Gradient Boosting	ST,IS	0.63	0.33	0.00	0.00	0.00	0.00
	ST,TH	0.42	0.32	0.00	0.00	0.00	0.00
	MT,IS	0.23	0.10	0.55	0.25	0.00	0.00
	MT,TH	0.00	0.00	0.62	0.51	0.00	0.00
	LT,IS	0.01	0.00	0.07	0.01	0.63	0.22
	LT,TH	0.00	0.00	0.00	0.00	0.17	0.21