On-the-fly learning of adaptive strategies with bandit algorithms

Rashid Bakirov, Bournemouth University, UK
Damien Fay, INFOR/Logicblox, Atlanta, GA, USA
Bogdan Gabrys, Advanced Analytics Institute, University of Technology Sydney, Ultimo, Australia

Introduction

The topic of this research is automatic adaptation strategies for stream learning.

- Non-stationary data streams require adaptation of the predictive mode
- Often multiple adaptive mechanisms (AM) available for each time step
- Which AM to deploy? A sequential decision problem.
- Here: Using multi armed bandit algorithms to solve this problem.
- Why? Computationally inexpensive, theoretical guarantees.
- Results comparable to custom algorithms.

Formulation

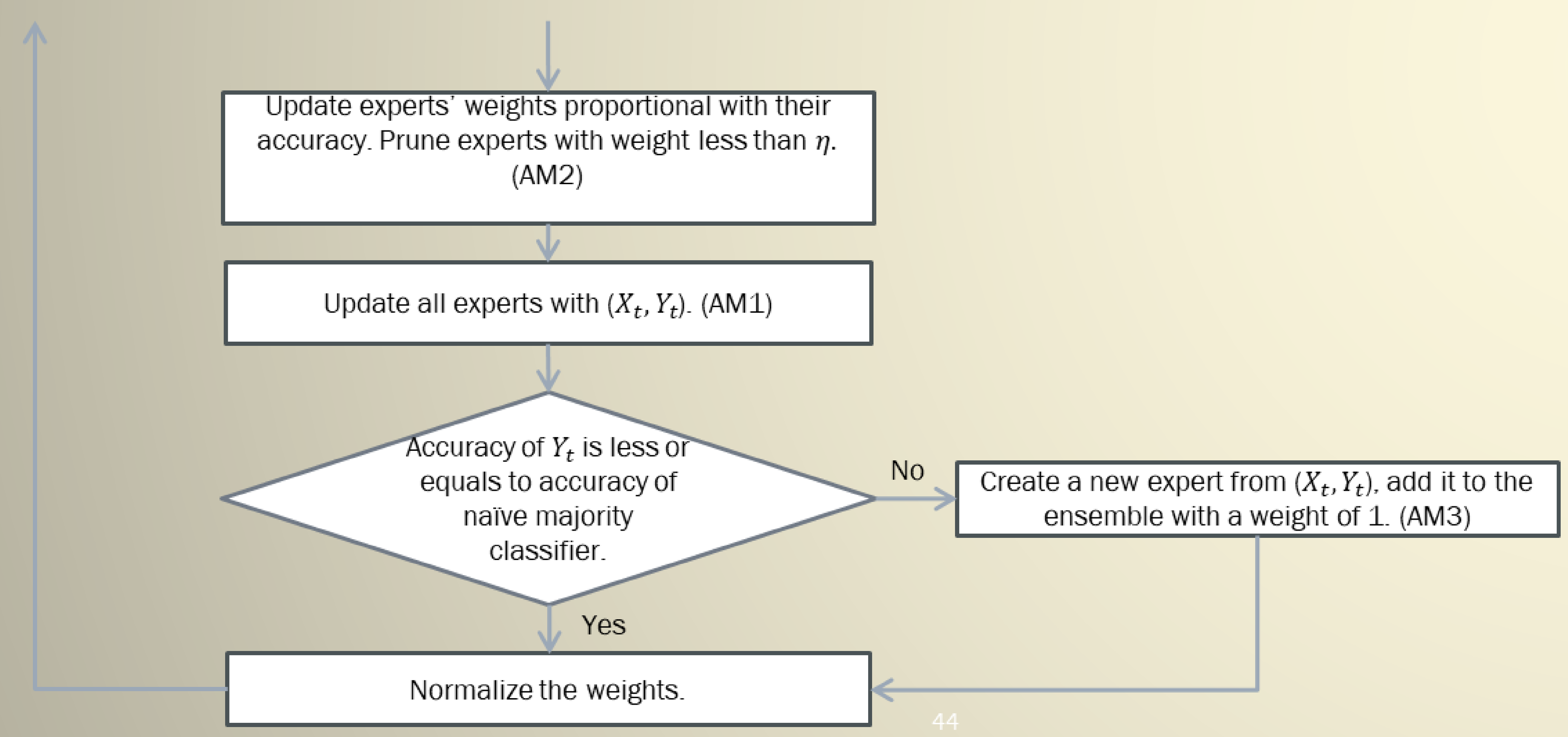
- Prediction: $\hat{y}_t = f_t(\boldsymbol{x}_t, \Theta_f)$
- After every time step true value y_t is revealed and loss $l_t(y_t, \hat{y}_t)$ is calculated
- Adaptive Mechanism: $g_t(f_t, \hat{y}, V_t, \Theta_g): f_t \rightarrow f_{t+1}$
- Multiple AMs: $\{\emptyset, g_1, \ldots, g_H\} = G$
- Adaptation: at each time step t, $f_{t+1} = g_{h_t}(f_t)$
- Adaptation sequence (AS):, $g_{h_1}, g_{h_2}, \ldots$
- Question: How to choose $h_1, h_2, \ldots$ to minimize the total loss $\sum_t l$ over $\{\boldsymbol{X}, \boldsymbol{y}\}$?

Empirical evaluation

- Dynamic Weighted Majority [Kolter and Maloof 2007]
- Paired Learner [Bach and Maloof 2010]

Comparing their *custom* (original) adaptation scheme to the ones learned by bandit algorithms.

- Cold vs "warm" start.
- All possible or only custom AMs.



Example: Dynamic Weighted Majority custom adaptation

Bandit algorithms

Comparing the *custom* (original) adaptation scheme to the ones learned by bandit algorithms:

- LinUCB [Li et al., 2010] – Contextual bandit algorithm
- KL-UCB [Garivier and Cappé, 2011] – Non-contextual algorithm
- ε-greedy algorithm - baseline

Context (for LinUCB)

- *Last deployed AM*
- *Last classification result*

Options

- Cold vs "warm" start.
- All possible or only custom AMs.

Data

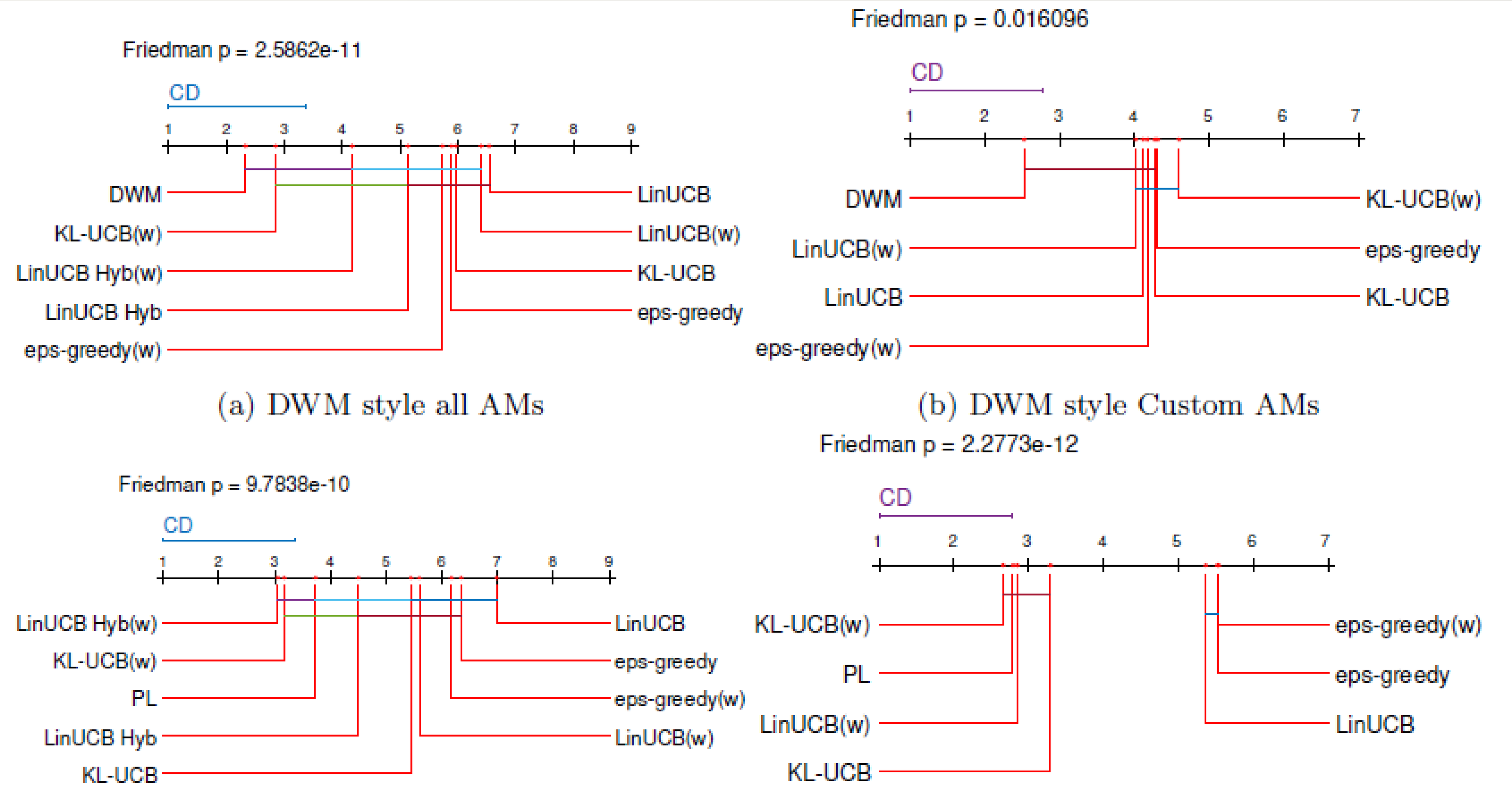
26 synthetic and 7 real-world classification data sets with different levels/types of non-stationarity

References

Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In Proceedings of the 19th International Conference on World Wide Web, WWW '10, page 661–670, New York, NY, USA, 2010. Association for Computing Machinery.

Aurélien Garivier and Olivier Cappé. The kl-ucb algorithm for bounded stochastic bandits and beyond. In Sham M. Kakade and Ulrike von Luxburg, editors, Proceedings of the 24th Annual Conference on Learning Theory, volume 19 of Proceedings of Machine Learning Research

Results



(a) DWM style all AMs
(b) DWM style Custom AMs
(c) PL style all AMs
(d) PL style custom AMs

Figure 1: Nemenyi plots (lower is better) of adaptation using bandit-based and custom adaptive strategies on synthetic datasets.

Table 1: Accuracy values of adaptation for bandit-based with all AMs and custom adaptive strategies on real datasets. Top accuracy values are bold, runner-ups are bold-italic.

a) DWM style	LinUCB	LinUCB(w)	LinUCB Hyb	LinUCB Hyb(w)	KL-UCB	KL-UCB(w)	ε-GRD	ε-GRD (w)	DWM
Electricity	0.855	0.843	0.853	0.853	0.845	***0.872***	0.847	0.847	**0.874**
Power	0.592	0.539	***0.675***	0.544	0.522	0.669	0.651	0.655	**0.677**
Contraceptive	0.418	***0.453***	0.412	0.381	0.379	**0.465**	0.379	0.388	0.399
Iris	0.748	0.864	0.808	0.846	0.866	0.873	***0.884***	0.849	**0.891**
Yeast	0.403	0.468	0.278	***0.468***	0.468	**0.469**	0.312	0.314	0.297
Gas	0.893	0.916	0.852	0.889	0.893	***0.925***	0.919	0.919	**0.943**
Gestures	0.917	0.915	0.911	0.923	0.913	***0.926***	0.883	0.881	**0.937**
b) PL style	LinUCB	LinUCB(w)	LinUCB Hyb	LinUCB Hyb(w)	KL-UCB	KL-UCB(w)	ε-GRD	ε-GRD (w)	PL
Electricity	0.860	***0.862***	0.861	0.859	0.859	0.858	0.839	0.840	**0.866**
Power	0.508	0.491	0.625	0.539	0.588	**0.678**	0.653	0.644	***0.645***
Contraceptive	**0.445**	***0.445***	0.445	0.434	0.434	0.433	0.410	0.415	0.413
Iris	0.730	0.858	0.856	0.860	***0.872***	0.733	0.870	0.828	**0.879**
Yeast	0.284	0.385	0.289	***0.470***	**0.470**	0.337	0.288	0.287	0.320
Gas	0.816	0.901	***0.910***	0.891	0.897	0.898	0.887	0.889	**0.929**
Gestures	0.848	0.849	0.841	***0.851***	0.847	0.845	0.775	0.773	**0.894**

Table 2: Accuracy values of adaptation for bandit-based with custom AMs and custom adaptive strategies on real datasets. Top accuracy values are bold, runner-ups are bold-italic.

a) DWM style	LinUCB	LinUCB(w)	KL-UCB	KL-UCB(w)	ε-GRD	ε-GRD(w)	DWM
Electricity	0.874	0.874	**0.875**	***0.875***	0.873	0.874	0.874
Power	0.673	**0.679**	***0.677***	0.675	0.676	0.675	0.677
Contraceptive	***0.403***	**0.407**	0.398	0.395	0.396	0.395	0.399
Iris	0.889	0.882	0.878	**0.893**	***0.891***	0.884	***0.891***
Yeast	***0.378***	**0.393**	0.330	0.312	0.325	0.364	0.297
Gas	***0.939***	0.938	0.937	0.936	0.939	0.939	**0.943**
Gestures	***0.930***	***0.930***	0.928	0.929	0.929	0.929	**0.937**
b) PL style	LinUCB	LinUCB(w)	KL-UCB	KL-UCB(w)	ε-GRD	ε-GRD(w)	PL
Electricity	0.860	0.861	***0.861***	0.860	0.860	0.859	**0.866**
Power	0.637	0.640	***0.670***	**0.675**	0.628	0.628	0.645
Contraceptive	0.409	0.409	0.408	0.409	**0.413**	0.408	**0.413**
Iris	**0.886**	0.867	0.853	0.856	0.860	0.860	***0.879***
Yeast	0.292	0.305	***0.332***	**0.333**	0.297	0.294	0.320
Gas	0.901	0.901	0.902	***0.907***	0.901	0.904	**0.929**
Gestures	0.849	0.842	***0.850***	***0.850***	0.843	0.846	**0.894**

Discussion

- Multiple Adaptive Mechanisms framework can adapt models on multiple levels, including the structure, parameters, hyperparameters, etc.
- **Adaptation strategies learned by bandit algorithms (LinUCB and KL-UCB) provide comparable empirical performance to custom adaptation algorithms, particularly with custom AMs and warm start.**
- LinUCB can provide theoretical guarantees in terms of suffered cumulative regret.
- Surprisingly, performance of LinUCB and KL-UCB is similar; hence the chosen context features need to be revisited.
- Full-information setting and "transfer learning" between different datasets for warm start can be investigated.