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

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.

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

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

Bandit algorithms
Comparing the custom (original) adaptation scheme to the ones learned by bandit algorithms:
- LinUCB [Li et al., 2010] – Contextual bandit algorithm
- $\epsilon$-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

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.

References