Towards AutoML in the presence of Drift

Jorge G. Madrid, Hugo J Escalante and Eduardo Morales*
INAOE, Puebla, México

Abstract

In many real-world scenarios AutoML systems are necessary due to the limited expertise in Machine Learning by developers. Research in this field has lead to the development of state-of-the-art solutions for supervised Machine Learning tasks that work effectively in different application areas. However, these don't consider the changing nature inherent to several real-world situations (e.g. recommendation systems, spam-filtering). We briefly describe a first attempt to cope with such situation. We extended Auto-Sklearn with intuitive mechanisms to allow it to cope with data distributions that change over time. Results demonstrate the effectiveness of the proposed methodology. An extended version of this work was presented in AutoML Workshop @ ICML2018.

1 Introduction

Autonomous Machine Learning (AutoML) is the field focusing on methods that aim at fully automating different stages of the machine learning process. Although progress in AutoML is vast, the considered scenarios are somewhat constrained, (e.g., in the type of approached problem, in the assumptions on data, in the size of datasets, etc). In this context, one of the most desirable features for AutoML methods is to work under a *lifelong machine learning* (LML) setting. LML refers to systems that can sequentially learn many tasks from one or more domains in its lifetime [11], these systems (not restricted to supervised learning) require the ability to retain knowledge, adapt to changes and transfer knowledge when learning a new task. An AutoML method that learns from different tasks and that is able to adapt itself during its lifetime would comprise a competitive and robust all-problem machine learning solution.

This report aims at exploring the viability of AutoML methods to operate in a LML setting, in particular in a scenario where the targets evolve over time, that is, in the presence of concept drift. We modify the Auto-Sklearn method with mechanisms that allow it to deal with the drift phenomenon in a simplified LML evaluation scenario. To the best of our knowledge these are the first results reported on AutoML in the presence of Concept Drift.

Although work for processing streams and data in the presence of drift is vast [10, 12, 3], to the best of our knowledge existing work as not approaching the AutoML setting: automatically building and updating a *full model* [1], that is a model that comprises data preprocessing, feature selection and classification model and that optimizes hyperparameters of the whole model.

2 Scenario, methods and experimental results

We consider a simplified LML - AutoML with a concept drift scenario, which is precisely the scenario considered in the forthcoming AutoML3[2] challenge. For the evaluation datasets are split in sequential batches, after each batch is predicted using the method to evaluate target values for that batch are revealed and the batch is evaluated. The overall performance is evaluated with the average of all the batches. Figure 1 illustrates the considered scenario.

*Authorship from the full-paper also includes: Wei Wei Tu, Yang Yu, Lisheng Sun-Hosoya Isabelle Guyon and Michele Sebag [8]

[2] `https://www.4paradigm.com/competition/nips2018`

1st LXAI Research Workshop at NIPS 2018. Montréal, Canada.

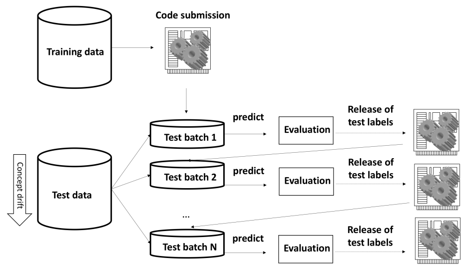

Figure 1: Evaluation scenario considered in the AutoML3 challenge.

We extended one of the most successful solutions in AutoML: Auto-Sklearn [2, 6, 5]. Our proposed method is firstly trained with a tagged batch of data, then it receives the following batch and makes predictions, after the target values are revealed it uses a explicit *Drift Detector* which rings an alarm if a drift in the data distribution occurred. Namely, we use FHDDM state-of-the-art method for drift detection in evolving data [9], FHDDM is based on the widely used method DDM[4] given its short number of hyper-parameters.

Instead of a single classification model, Auto-Sklearn generates an ensemble which our method *improves* after drift is detected, we describe very shortly the 4 strategies we propose in order to achieve this.

- **Model replacement.** A traditional concept drift adaptation where the model is globally replaced with a new one.
- **WU-latest.** Ensemble weights are updated with the latest data batch
- **WU-all.** Ensemble weights are updated with all batches processed.
- **Add new.** A new set of models are trained and added to the current ensemble.

We tested the 4 different strategies in benchmark Concept Drift datasets and compared them with the original Auto-Sklearn method. Experiments with undisclosed AutoML data [7] with presence of temporal dependencies were also performed. As shown in the tables, the 4 strategies of our method improve the baseline. The best performing strategy depends mainly on the type of drift present in the dataset (gradual, abrupt, recurrent, etc.) and the overall structure of its features.

Figure 2: Accuracy per batch of our methods. From left to right: Electricity, Poker and Chess datasets.

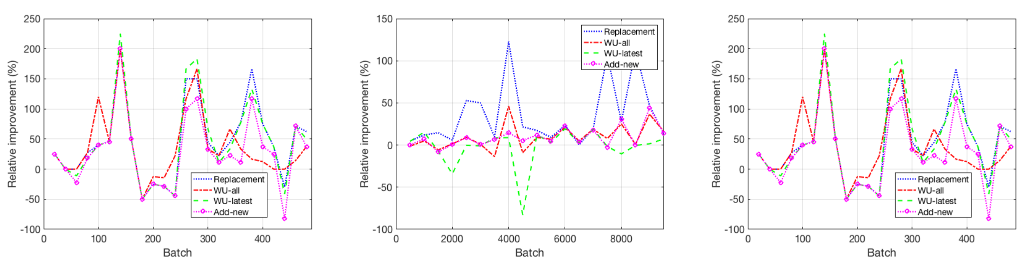

Table 1: Results on data from the AutoML2 challenge (AUC).

Method	PM*	RH	RI	RL	RM*
Base	0.433	0.192	0.299	0.340	0.264
Replacement	0.433	0.197	0.092	**0.478**	0.264
WU-all	0.433	**0.370**	0.199	0.212	0.264
WU-latest	0.433	0.270	**0.450**	0.405	0.264
Add new	0.433	0.298	0.184	0.277	0.264

3 Conclusions

To the best of our knowledge this is the first work dealing with AutoML in the presence of drift. Experimental results confirm the usefulness of the proposed mechanisms for dealing with drift. Although our results are far from being conclusive, they bring some light into the performance of AutoML for evolving data streams.

References

[1] Hugo Jair Escalante, Manuel Montes, and Luis Enrique Sucar. Particle swarm model selection. *J. Mach. Learn. Res.*, 10:405–440, June 2009.

[2] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. Efficient and robust automated machine learning. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 2962–2970. Curran Associates, Inc., 2015.

[3] João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. A survey on concept drift adaptation. *ACM Comput. Surv.*, 46(4):44:1–44:37, March 2014.

[4] João Gama, Pedro Medas, Gladys Castillo, and Pedro Rodrigues. Learning with drift detection. In Ana L. C. Bazzan and Sofiane Labidi, editors, *Advances in Artificial Intelligence – SBIA 2004*, pages 286–295, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[5] Isabelle Guyon, Kristin Bennett, Gavin C. Cawley, Hugo Jair Escalante, Sergio Escalera, Tin Kam Ho, Núria Macià, Bisakha Ray, Mehreen Saeed, Alexander R. Statnikov, and Evelyne Viegas. Design of the 2015 chalearn automl challenge. In *2015 International Joint Conference on Neural Networks, IJCNN 2015, Killarney, Ireland, July 12-17, 2015* [5], pages 1–8.

[6] Isabelle Guyon, Imad Chaabane, Hugo Jair Escalante, Sergio Escalera, Damir Jajetic, James Robert Lloyd, Núria Macià, Bisakha Ray, Lukasz Romaszko, Michèle Sebag, Alexander R. Statnikov, Sébastien Treguer, and Evelyne Viegas. A brief review of the chalearn automl challenge: Any-time any-dataset learning without human intervention. In *Proceedings of the 2016 Workshop on Automatic Machine Learning, AutoML 2016, co-located with 33rd International Conference on Machine Learning (ICML 2016), New York City, NY, USA, June 24, 2016*, pages 21–30, 2016.

[7] Isabelle Guyon, Lisheng Sun-Hosoya, Marc Boullé, Hugo Jair Escalante, Sergio Escalera, Zhengying Liu, Damir Jajetic, Bisakha Ray, Mehreen Saeed, Michele Sebag, Alexander Statnikov, Wei-Wei Tu, and Evelyne Viegas. Analysis of the automl challenge series 2015-2018. In Frank Hutter et al., editor, *AutoML, Springer Series on Challenges in Machine Learning (to appear)*. 2018.

[8] Jorge G Madrid, Hugo J Escalante, Eduardo Morales, Wei Wei Tu, Yang Yu, Isabelle Guyon Lisheng Sun-Hosoya, and Michele Sebag. Towards automl in the presence of drift: first results. In *Workshop on AutoML 2018, at the 35th International Conference on Machine Learning (ICML 2018), Stockholm, Sweden, July 14, 2018.*

[9] Ali Pesaranghader, Herna Viktor, and Eric Paquet. Reservoir of diverse adaptive learners and stacking fast hoeffding drift detection methods for evolving data streams. *CoRR*, abs/1709.02457, 2017.

[10] Jeffrey C Schlimmer and Richard H Granger. Beyond incremental processing: Tracking concept drift. In *AAAI*, pages 502–507, 1986.

[11] Daniel L. Silver and Qiang Yang amd Lianghao Li. Lifelong machine learning systems: Beyond learning algorithms. In *Lifelong Machine Learning: Papers from the 2013 AAAI Spring Symposium*, pages 49–55, 2013.

[12] Gerhard Widmer and Miroslav Kubat. Learning in the presence of concept drift and hidden contexts. *Machine learning*, 23(1):69–101, 1996.