Just-in-Time Adaptive Classifiers

Cesare Alippi, Manuel Roveri
DEI, Politecnico di Milano, Piazza L. Da Vinci 32, 20133 Milano, Italy
Alippi, Roveri@elet.polimi.it

Many real world systems are characterized by evolving dynamics and may change their behaviour over lifetime. In general, this is due to ageing effects, thermal drifts, soft and hard faults that affect industrial, environmental and natural phenomena and, as a consequence, the data generation process. In solutions developed for such applications, e.g., quality assessment and control, environmental monitoring, the process evolution requires adaptive mechanisms for tracking the system evolution in order to maintain acceptable performance.
Adaptive solutions envisaging classification systems e.g., [1][2], generally assume that the process generating the data is stationary, hypothesis hard to be satisfied in many real applications; on the non-stationary front there it exists a limited literature addressing the design of classification systems, e.g., see [3-5].

Here, we focus on adaptive classifiers which, by *exploiting information coming from the field during operational life* (and constituting an incremental knowledge base), modify whenever appropriate their knowledge base to maintain/improve classification accuracy. In particular, the classifier integrates fresh available information in stationary conditions and reacts to the changing environment in non-stationary ones to track the system change. The novelty of the approach resides in the possibility to update just-in-time, i.e., exactly when needed or convenient, the classifier. Design of a just-in-time adaptive classifier involves two steps:

1. definition of methods estimating the most convenient instant of time for integrating the incremental knowledge base in the classifier (stationary case, e.g., through sequential analysis) or associated with a loss in stationarity (non-stationary case, e.g., see [6]);
2. integration of the incremental knowledge base in the knowledge base of the classifier to track the process change.

In stationary situations adaptation, by integrating novel information into the knowledge base of the classifier, improves the classifier performance. Likewise, in non-stationary conditions, the adaptive classifier operates by removing, when needed, obsolete information and integrating fresh ones in the knowledge base so as to track the system evolution.

Just-in-time adaptive classifiers in stationary conditions: improving the classification accuracy
Integration of novel information in adaptive classifiers during operational life is profitable in hierarchical classification families supporting consistent rules, e.g., neural networks and k-NN classifiers. Particularly appealing is the use of k-NN classifiers for their peculiar "computation-free" training phase and the easy information management modality. In k-NNs selection of the optimal classifier complexity k as the number of available samples n increases is generally carried out by means of a trial-and-error approach (e.g., by relying on cross-validation or Leave-One-Out –LOO- techniques). A LOO estimate of k does not require any a priori information and, when n asymptotically increases, tends to its optimal value; unfortunately, its evaluation is generally computationally expensive. Differently, Fukunaga [7] tackles the problem of identifying the optimal k from a theoretical point of view. There it is required a priori information such as the pdf of the process generating the data, information which is not always available. We suggest, by integrating LOO and Fukunaga approaches, an estimate of the optimal k that is easy to use and does not require a priori information. In stationary conditions and under the asymptotical assumption (e.g., $n > 50$) k can be evaluated as,

$$\hat{k} = \hat{k}_{LOO}^{0} \left(\frac{n}{n_0} \right)^{\frac{4}{d+4}} \tag{1}$$

where d is the size of the feature space, $\hat{k}_{LOO}^{0}$ is the estimate of k provided by LOO on the initial knowledge base, n is the current number of samples stored in the knowledge base and n_0 is the number of samples in the initial knowledge base. Introduction of new samples in the knowledge base of the classifier implies $\hat{k}$ to increase.

Just-in-time adaptive classifiers in non-stationary conditions: tracking the process change
The joint use of the process change detection tests and the dynamic knowledge management allow us for defining a methodology able to design just-in-time adaptive classifiers working in non-stationary environments; the proposed methodology is provided in Algorithm 1.

The initial knowledge base KB_0 characterizing the initial k-NN classifier is used to estimate $\hat{k}^0_{LOO}$ (Step 1), then configure the classifier (Step 2) and the non-stationary change detection test (Step 3). When new information stored in the incremental knowledge base *IKB* is provided to the classifier (Step 7), it is suitably integrated in the knowledge base *KB* (Step 8) and the new k is computed (Step 10). If the process under monitoring remains stationary and new knowledge is not available during operational life, the methodology works as in traditional classifiers (Step 14). When the change detection test detects a non-stationary behavior for the process (Step 15), obsolete information is removed from the knowledge base (Step 17), a re-configuration of the classifier is invoked (Step 20) and re-configuration of the non-stationary change detection test is performed (the new process becomes the reference one for detecting subsequent evolutions – Step 21-).

Algorithm 1: Just-in-time Adaptive Classifiers

1. Estimate $\hat{k}^0_{LOO}$ by means of LOO applied to KB_0;
2. Configure the classifier on KB_0;
3. Configure the non-stationarity detection test on KB_0 ;
4. $KB(0)=KB_0$;
5. $i=1$;
6. **while** (1) **do**
7.   **if** (new knowledge *IKB(i)* is available) **then**
8.   *KB(i)=Add(KB(i-1),IKB(i))*;
9. $n = |KB(i)|$;
10. $\hat{k} = \hat{k}^0_{LOO}\left(\frac{n}{n_0}\right)^{\frac{4}{d+4}}$
11. $i=i+1$;
12.   **end if**
13.   **if** (non-stationarity detection test (sample *x)* == Stationary) **then**
14.   classification= k-NN(x, $\hat{k}$)
15.   **else**
16. KB_{OBS}= old knowledge (more than last *T* samples)
17. $KB(0)=Rem(KB(i), KB_{OBS})$;
18. $n_0 = |KB(0)|$;
19.   Estimate $\hat{k}^0_{LOO}$ on *KB(0)*
20.   Configure the classifier on *KB(0)*
21.   Configure the non-stationarity detection test on *KB(0)*;
22. $i=1$;
23.   **end if**
24. **end while**

A large experimental campaign addressing both industrial and synthetic datasets shows the advantages of just-in-time adaptive classifiers both in stationary, non-stationary and mixed environments. Results, together with a critical view of the methodology and details will be presented at the workshop.

References:

[1] K. Nozaki, H. Ishibuchi, and H. Tanaka, “Adaptive fuzzy rule-based classification systems”, Fuzzy Systems, IEEE Transactions on, vol. 4, no. 3, pp. 23-250, Aug. 1996

[2] M.M. Rizki, M.A. Zmuda, and L.A. Tamburino, “Evolving pattern recognition systems”, Evolutionary Computation, IEEE Transactions on, vol. 6, no. 6, pp. 594-609, Dec. 2002

[3] Jianqi Wang, M.R. Azimi-Sadjadi, and D. Reinke, “A temporally adaptive classifier for multispectral imagery”, Neural Networks, IEEE Transactions on, vol. 15, no. 1, pp. 159-165, Jan. 2004

[4] L. Rutkowski, “Adaptive probabilistic neural networks for pattern classification in time-varying environment”, IEEE Trans. on Neural Networks, vol. 15, pp. 811-827, 2004

[5] G.A. Carpenter, S. Grossberg, and J.H. Reynolds, “A fuzzy artmap nonparametric probability estimator for nonstationary pattern recognition problems”, Neural Networks, IEEE Transactions on, vol. 6, no. 6, pp. 1330-1336, Nov. 1995

[6] C. Alippi and M. Roveri, ”Process Drifts and Non-stationarity detecting indexes”, Proc. Learning, Snowbird-Utah, April 2006

[7] K. Fukunaga and L. Hostetler, “Optimization of k nearest neighbor density estimates,” Information Theory, IEEE Transactions on, vol. 19, no. 3, pp. 320-326, 1973