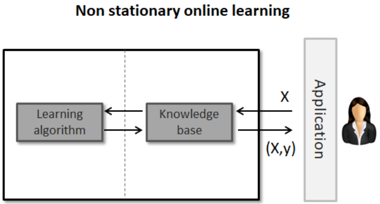
Drift anticipation with forgetting to improve evolving fuzzy system

Clément Leroy, Éric Anquetil, Nathalie Girard
clement.leroy@irisa.fr

Univ Rennes, CNRS, IRISA - Intuidoc Team
www-intuidoc.irisa.fr

Outline

Datastream learning context



- Command/Gesture recognition system
- Recommendation system
- Fraud detection

Datastream characteristics	Learning model requirements
- Continuous data	- Continuous adaptation - Fast adaptation - Prediction anytime
- Potentially infinite data	- Fix the memory space -> one-shot learning
- Non-stationary stream	- Adapt to different kind of concept drifts (incremental, brutal ...)

Outline

Research framework

Related works

Our proposal: Anticipate forgetting in conclusion part

Experimental validation

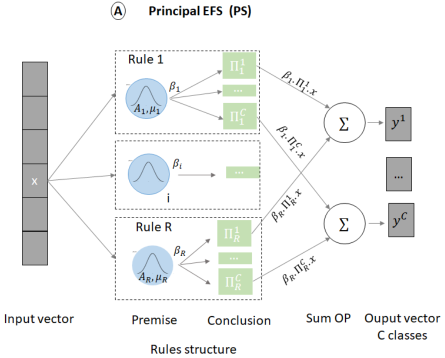
IF *Premise* **THEN** *Conclusion*

TS Rule i: IF $\mathbf{x}$ is close to P_i **THEN** $y_i^1 = \Pi_i^1 \mathbf{x} \; .. \; y_i^k = \Pi_i^k \mathbf{x}$

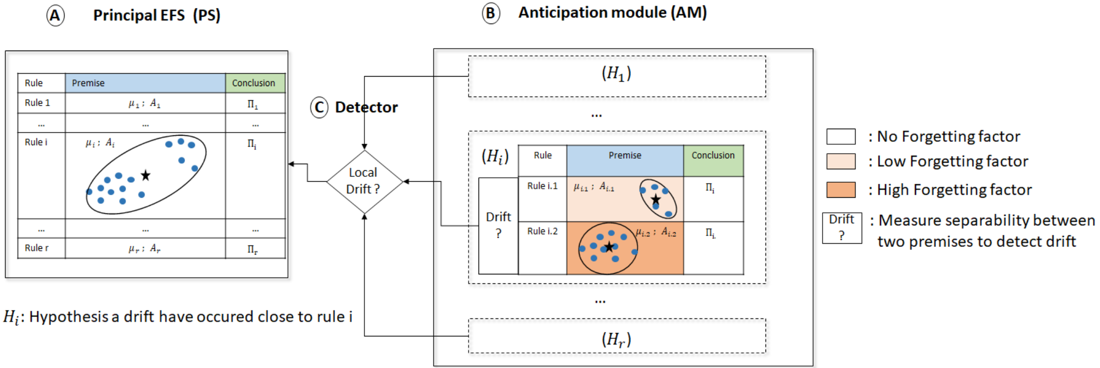
$\beta \in [0, 1]$: membership function

$$y^m = \sum_{i=1}^{r} \beta_i(\mathbf{x}) y_i^m$$

$$class(\mathbf{x}) = y = argmax \; y^m(\mathbf{x})$$

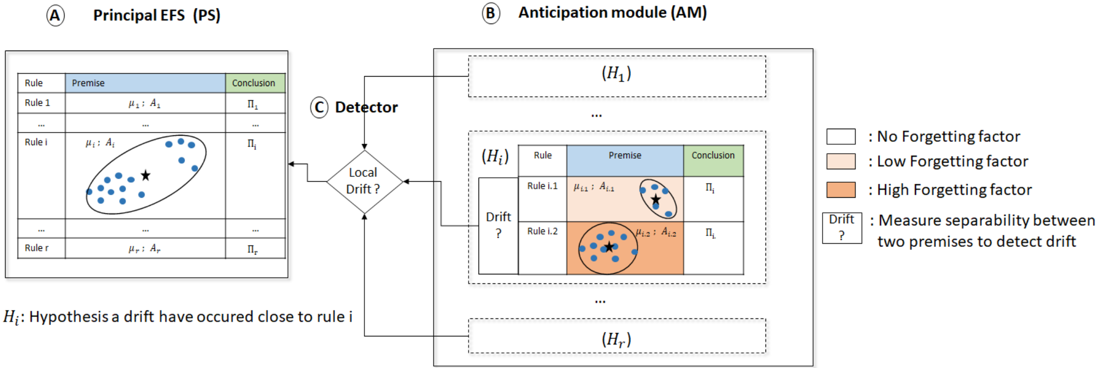


ParaFIS: A design based on the anticipation concept



[Leroy 2019]

ParaFIS: A design based on the anticipation concept



[Leroy 2019]

+ More reactive than starting from scratch
+ More stable than always forget

Outline

Further in the anticipation concept

What about anticipation in conclusion?

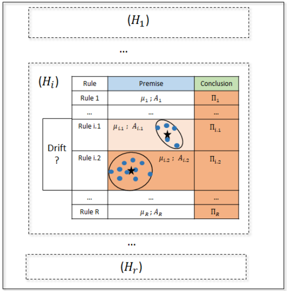
+ Integrate forgetting capacity in conclusion (improve reactivity)
+ Restoring consistency between premise part and conclusion part

Two main difficulties to deal with:

- Forgetting in the weighted recursive least square learning model (WRLS) not fully solve
    - Windup problem, collapse of conclusion, noise sensitivity
- Interconnection between premises and conclusions
    - weight: $\bar{\beta}_i(x) = \frac{\beta_i(x)}{\sum_j \beta_j(x)}$

$\rightarrow$ Complexity in R*(R+1)*C $\Longrightarrow$ Real-time constraint not satisfied

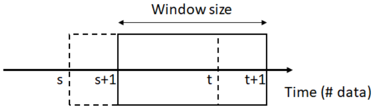
(B) **Anticipation module (AM)**



The differed directional forgetting - DDF

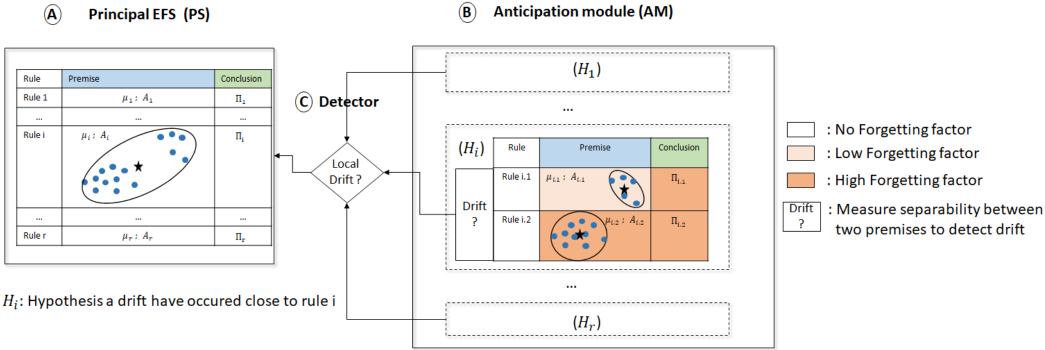
Idea: Apply DDF **[Bouillon 2012]** only in the anticipation module

Forgetting factor = Correlation matrix
Compute on a sliding window
→ Forget in the direction of data activation



+ Preserve stability in the principal system - no windup problem

Reduce complexity

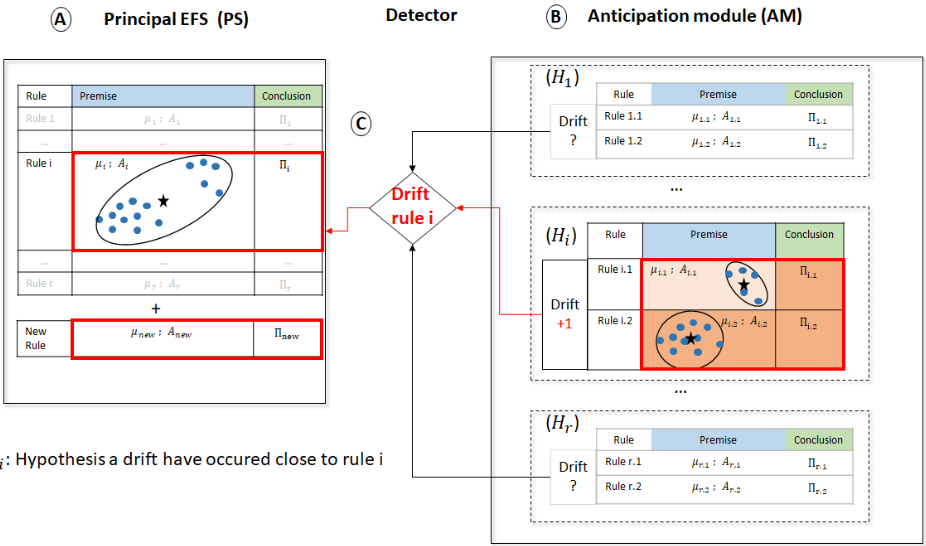


Reduce the number of parameters to learn

Two strategies to replace conclusion after a drift:

- The Naive Approach
- The Global Approach

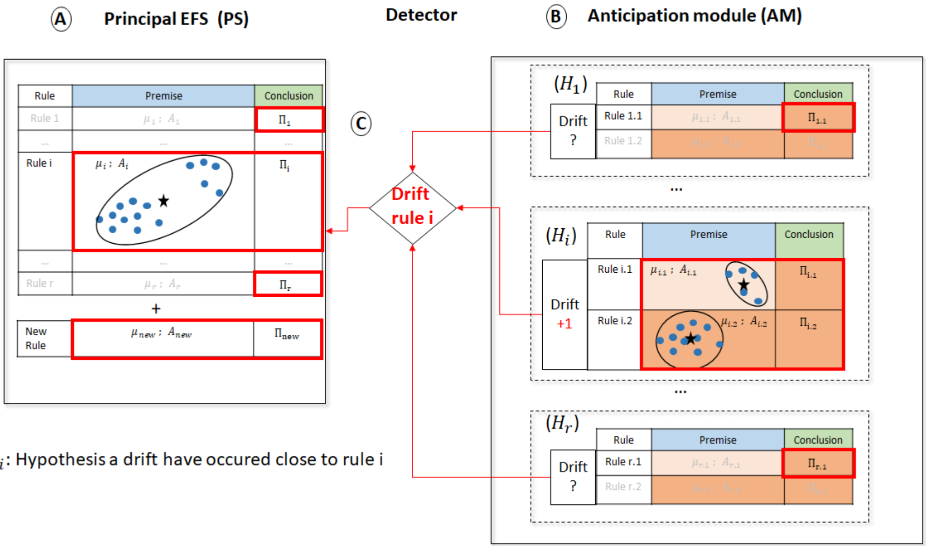
Reduce complexity: Naive Approach



H_i: Hypothesis a drift have occured close to rule i

Naive Approach !

Reduce complexity: Global Approach

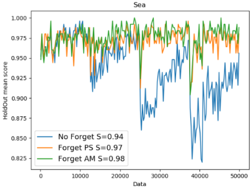


H_i: Hypothesis a drift have occured close to rule i

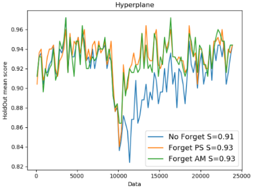
Global Approach !

Outline

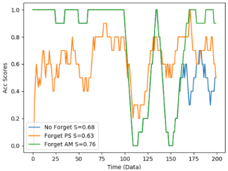
Experimental validation



(a) Sea



(b) Hyperplane



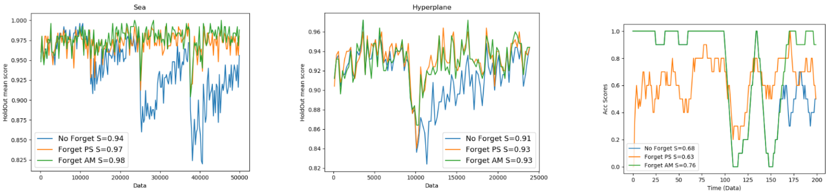
(c) 10dplane

TABLE I: Final Results - Mean Accuracy Score

Model		Electricity Pricing	Hyperplane	Iris+	Car	10dplane	Weather	Sea	SinH	Line	Sin	Mean
ParaFIS	No Forget	**77±15**	91±03	82±14	8±11	68±34	78±03	94±04	67±09	85±15	85±13	81
	Forget PS	**77±15**	**93±02**	**85±12**	79±12	63±16	78±03	97±01	71±07	93±06	**94±06**	83
	Forget AM Naive	**77±15**	**93±02**	82±14	**82±09**	70±31	79±03	96±03	70±07	92±10	93±09	83
	Forget AM Global	**77±15**	**93±02**	**85±15**	81±10	77±34	78±03	**98±01**	70±07	**94±06**	**94±08**	**85**

The level of significance of these results have been validated by a McNemar test

Experimental validation



(a) Sea (b) Hyperplane (c) 10dplane

TABLE I: Final Results - Mean Accuracy Score

Model		Electricity Pricing	Hyperplane	Iris+	Car	10dplane	Weather	Sea	SinH	Line	Sin	Mean
ParaFIS	No Forget	**77±15**	91±03	82±14	8±11	68±34	78±03	94±04	67±09	85±15	85±13	81
	Forget PS	**77±15**	**93±02**	**85±12**	79±12	63±16	78±03	97±01	71±07	93±06	**94±06**	83
	Forget AM Naive	**77±15**	**93±02**	82±14	**82±09**	70±31	79±03	96±03	70±07	92±10	93±09	83
	Forget AM Global	**77±15**	**93±02**	**85±15**	81±10	77±34	78±03	**98±01**	70±07	**94±06**	**94±08**	**85**
Learn++	CDE	69±08	90±00	**85±14**	68±30	71±13	73±02	93±02	**75±50**	89±14	80±13	79
Learn++	NSE	69±08	91±02	84±17	67±30	72±14	75±03	93±02	73±22	88±13	80±15	79
pENsemble	AxisParallel	75±16	92±02	78±15	79±10	78±20	**80±02**	97±02	71±06	90±07	78±26	82
pENsemble	Multivariate	75±16	92±02	75±17	79±10	**80±20**	78±02	97±02	71±06	90±07	78±30	82
pClass		68±10	91±02	73±18	77±10	63±26	68±04	89±10	71±09	91±07	72±20	76