POLITECNICO
MILANO 1863

Addressing Temporal Dependence, Concept Drifts, and Forgetting in Data Streams

Federico Giannini

Joint work with: Emanuele Della Valle

DEIB - Department of Electronics, Information Science, and Bioengineering

Politecnico di Milano, Milano, Italy

DELTA Workshop at ACM SIGKDD International Conference on Knowledge Discovery and Data Mining

26th August 2024, Barcelona, Catalonia (Spain)

1) Challenges of Learning from Data Streams

	SML	CL	TSA
1) Managing changes in data distribution (**concept drifts**).	X	X	
2) Learning continuously from single data points or mini-batches.	X		
3) Remembering all the acquired knowledge (avoid **catastrophic forgetting**).		X	
4) Handling **temporal dependence**.			X

- **SML**: Streaming Machine Learning
- **CL**: Continual Learning
- **TSA**: Time Series Analytics

Ziffer, G. et al. **Towards Time-Evolving Analytics: Online Learning for Time-dependent Evolving Data Streams.** Data Science (Preprint), 1–16

2) Research Goal

How do streaming models behave in the case of elaborated temporal dependence?

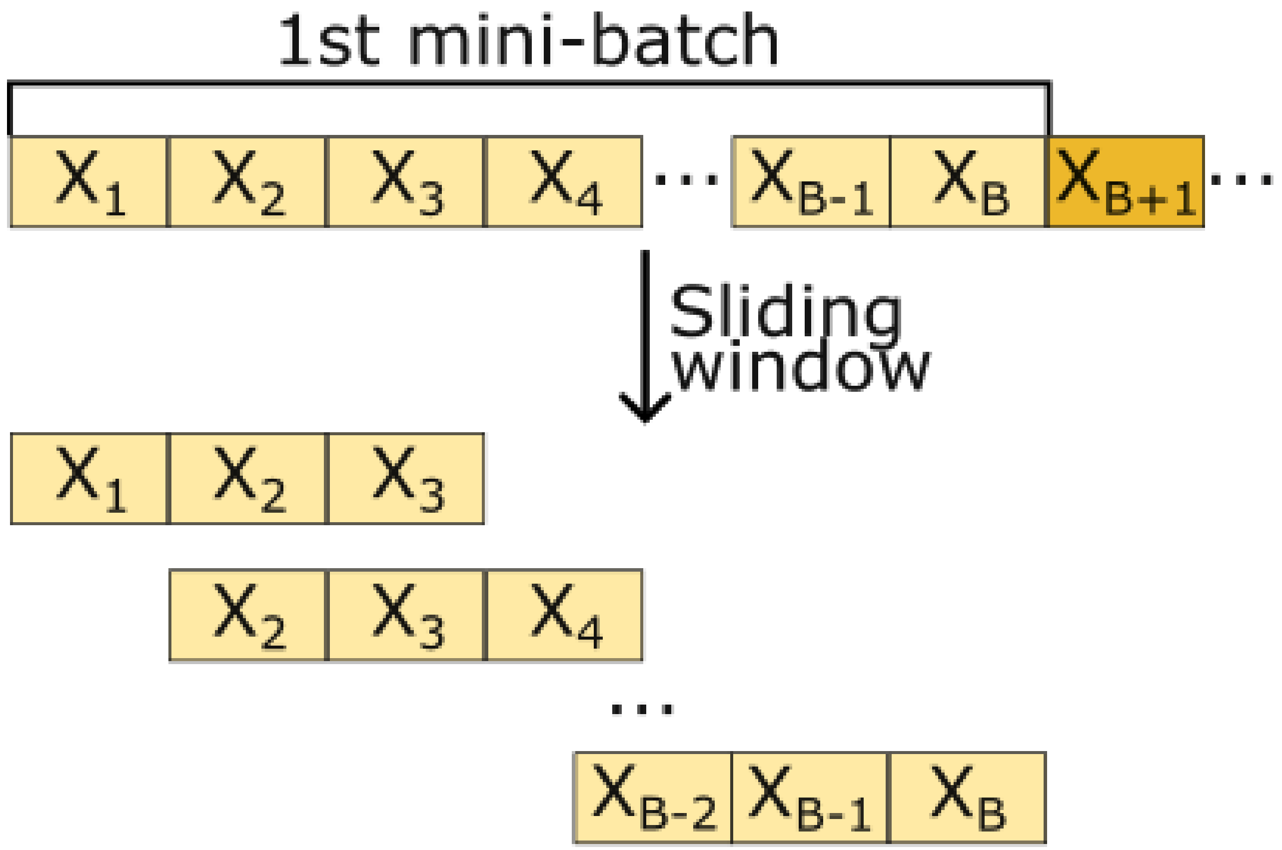
Comparison between:

- SML models
- SML models with **temporal augmentation**

$$X_t^{TA} = X_t \cup \{y_{t-1}, \dots, y_{t-k}\}$$

- Our previous solution **cPNN**

3a) cPNN: Taming Temporal Dependence when Learning Continuously



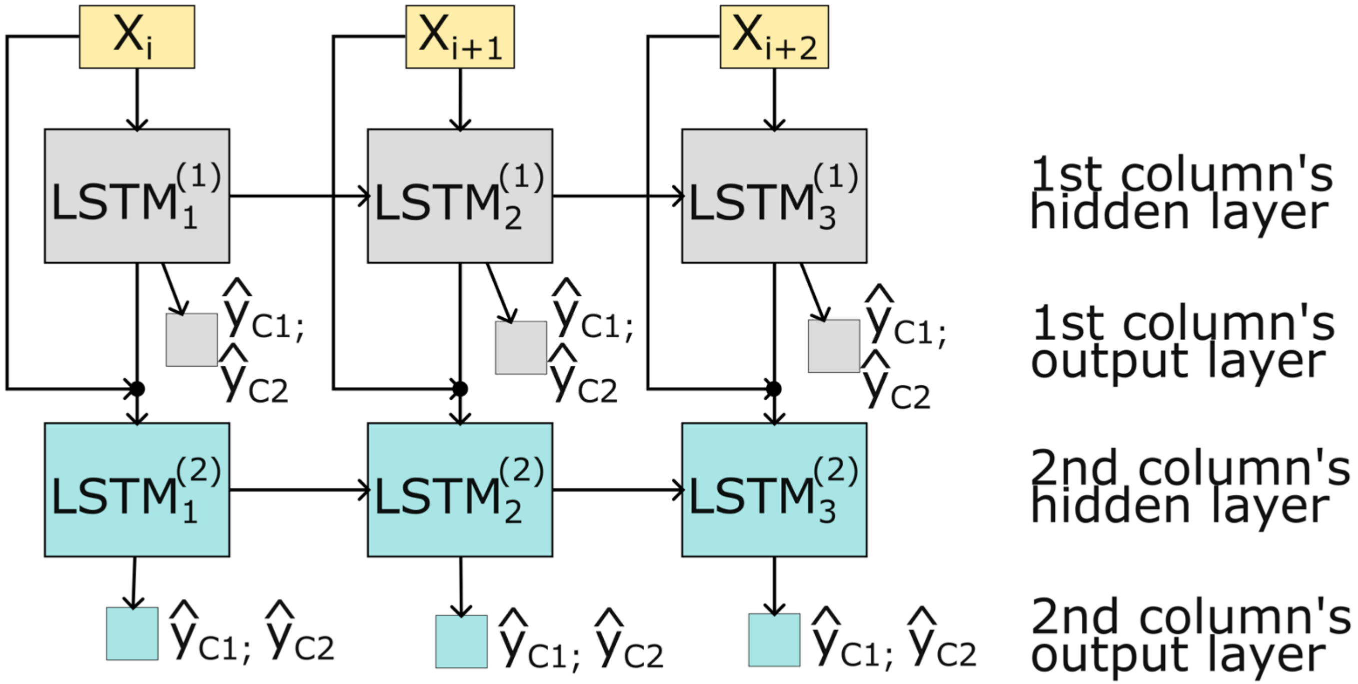
- Accumulate data points in **mini-batches.**
- Apply **windowing** on the mini-batch to build the **sequences.**
- Input the sequences to LSTM (**cLSTM**).

Giannini, F., Ziffer, G., & Della Valle, E. (2023). **cPNN: Continuous Progressive Neural Networks for Evolving Streaming Time Series**. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 328-340).

Lemos Neto, Á. C., Coelho, R. A., & Castro, C. L. D. (2022). **An Incremental Learning Approach Using Long Short-term Memory Neural Networks**. *Journal of Control, Automation and Electrical Systems*, *33*(5), 1457-1465.

3b) cPNN: Addressing Concept Drifts and Forgetting

- **Transfer learning** to adapt to new concepts quickly.
- **Freezing the weights** to avoid catastrophic forgetting.



https://github.com/federicogiannini13/cPNN_extended

Rusu, A.A et al. **Progressive Neural Networks**. CoRR abs/1606.04671 (2016)

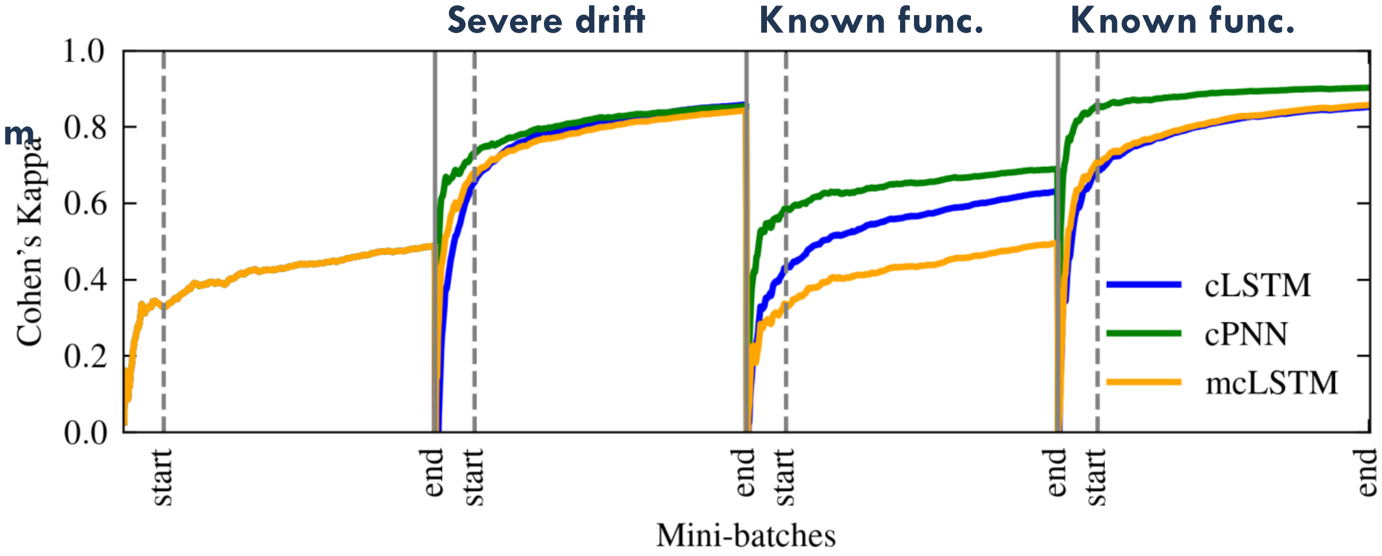
Cossu, A. et al. **Continual Learning with Gated Incremental Memories for Sequential Data Processing**. In: IJCNN. pp. 1–8. IEEE (2020)

4a) Experiments: Setting

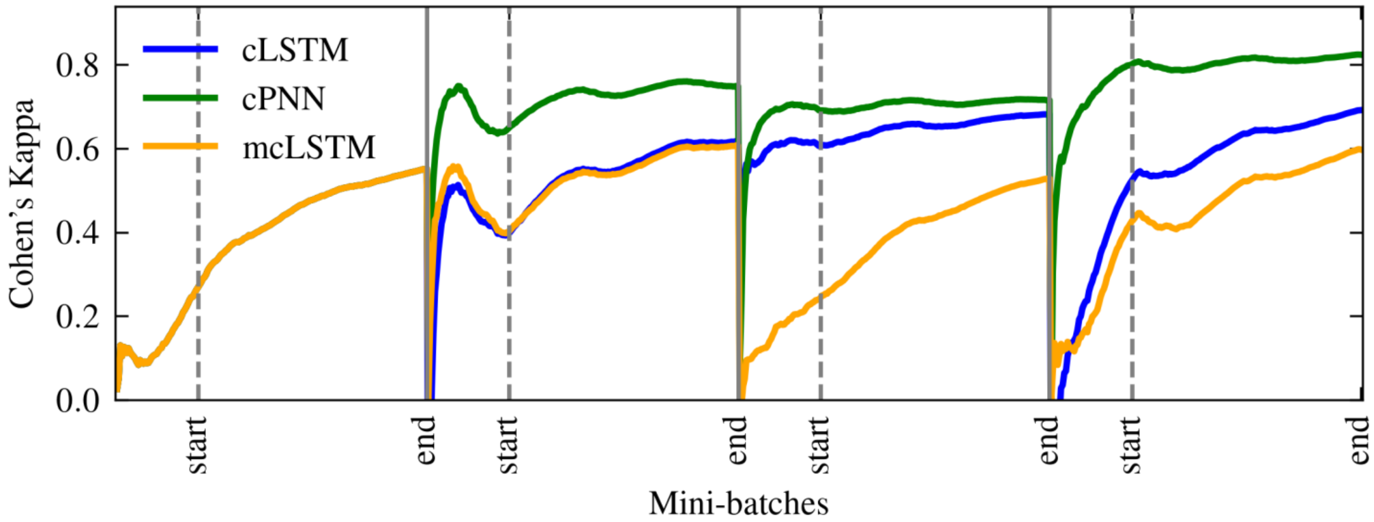
- Both **synthetic** and **real** data streams.
- Three **abrupt concept drifts**.
- Some drifts are **severe** (they change more than 50% labels).
- The remaining drifts are called **mild**.
- Some boundary functions are reproposed.

4b) Experiments: Ablation Study

Synthetic data stream



Weather data stream



Ablated versions:

- cLSTM: no drift management
- mcLSTM: no transfer learning

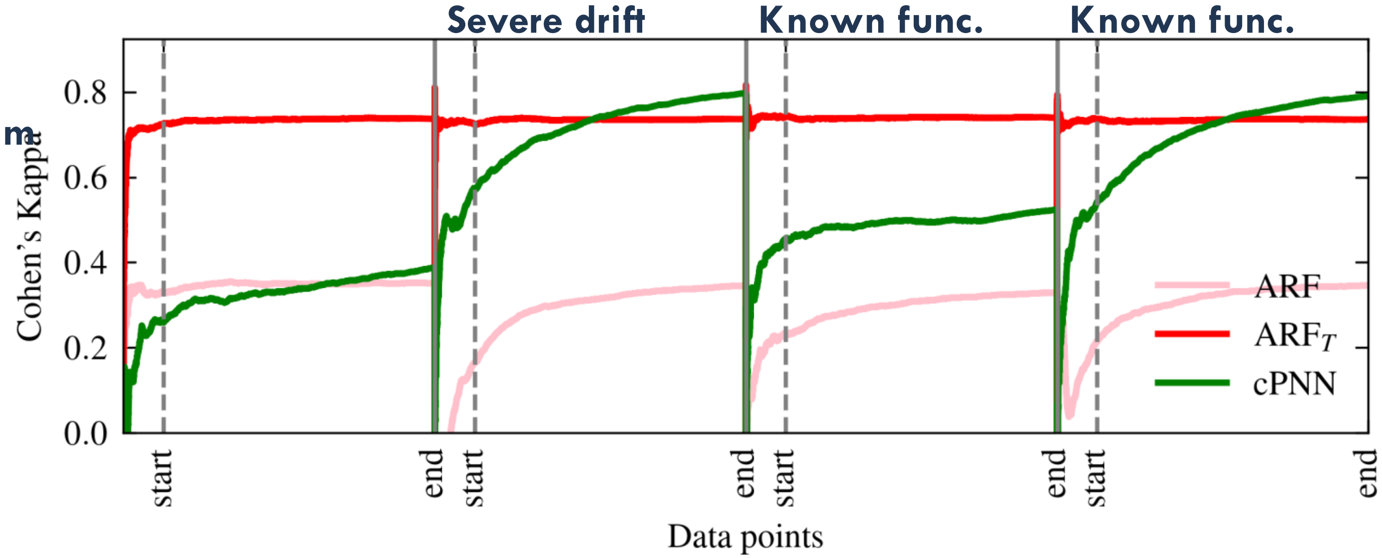
cPNN adapts quicker.

When the drift is severe cLSTM struggles.

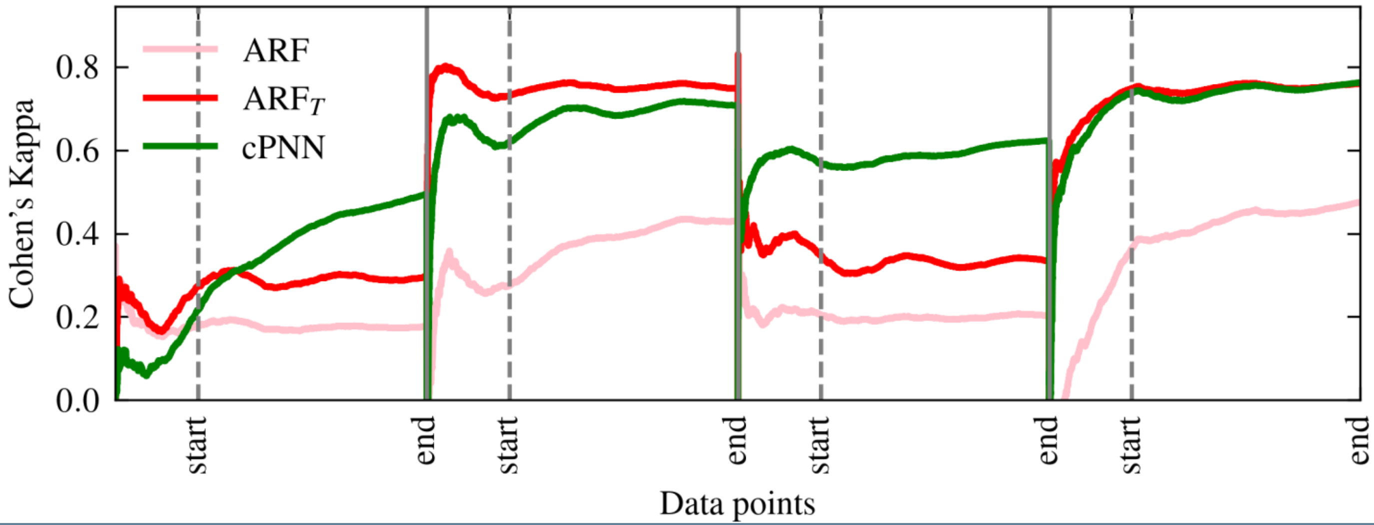
mcLSTM is quite always the worst.

4c) Experiments: Comparison with SML Models

Synthetic data stream



Weather data stream



ARF model cannot learn.

Temporal augmentation has a bias on the previous label.

cPNN can learn properly all the functions.

5) Conclusions and Future Works

Conclusions

- cPNN pioneers a solution to deal jointly with all the challenges of data streams.
- cPNN adapts quickly to concept drifts thanks to transfer learning.
- SML models cannot learn with temporal dependence.
- Temporal augmentation is biased on the previous label.

Future works

- Quantisation to reduce the memory footprint.
- Study of forgetting.
- Federated Learning.

Thank you!

Any questions?

Federico Giannini

federicogiannini13

federico.giannini@polimi.it

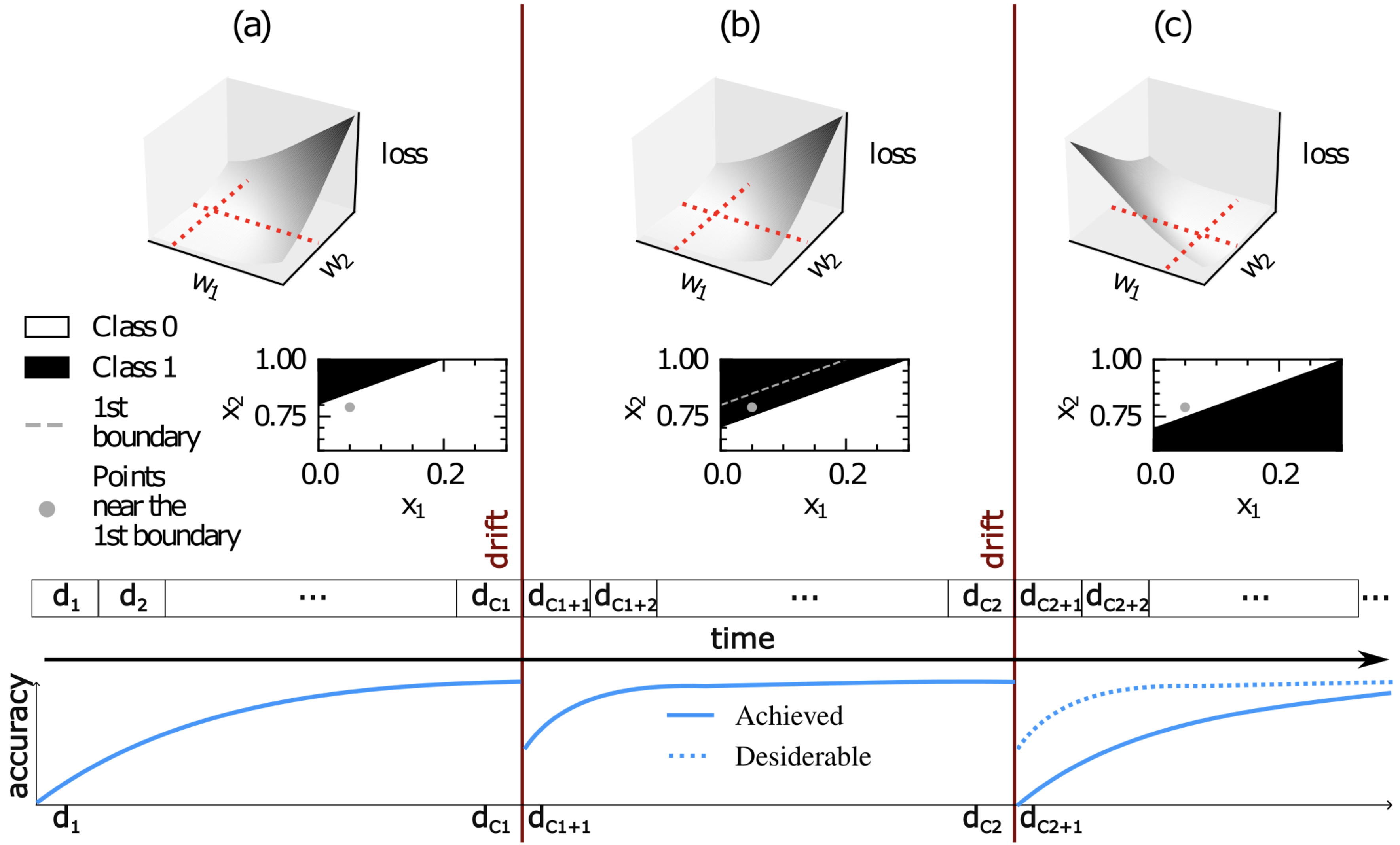
https://github.com/federicogiannini13/cPNN_extended

Extra Slides

Outline

1. Challenges of learning in dynamic environments
2. Research goal
3. cPNN: Continuous Progressive Neural Networks
4. Experiments
5. Conclusions and Future Works

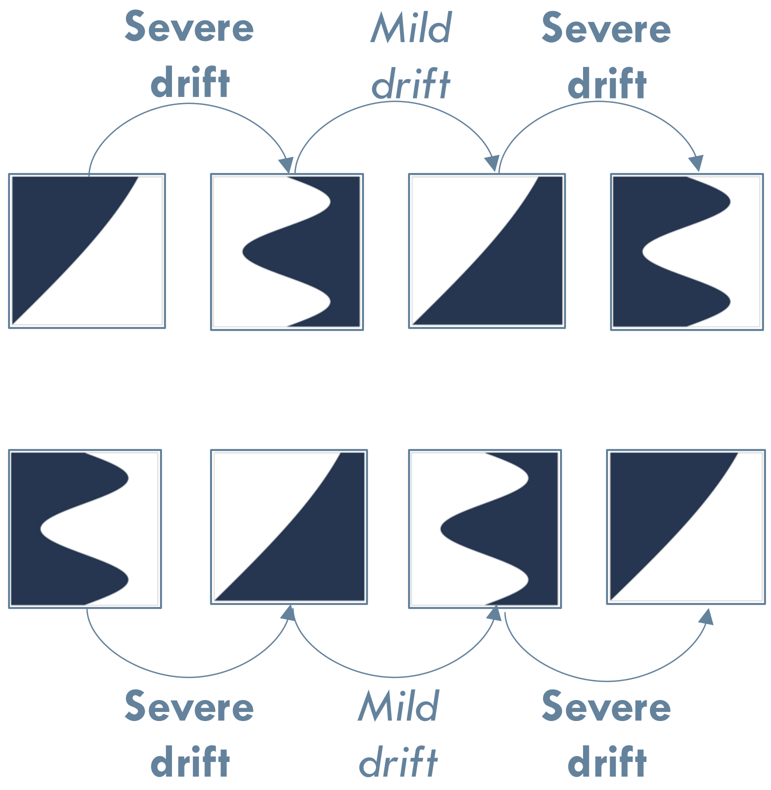
Problems of SGD for Evolving Data Streams



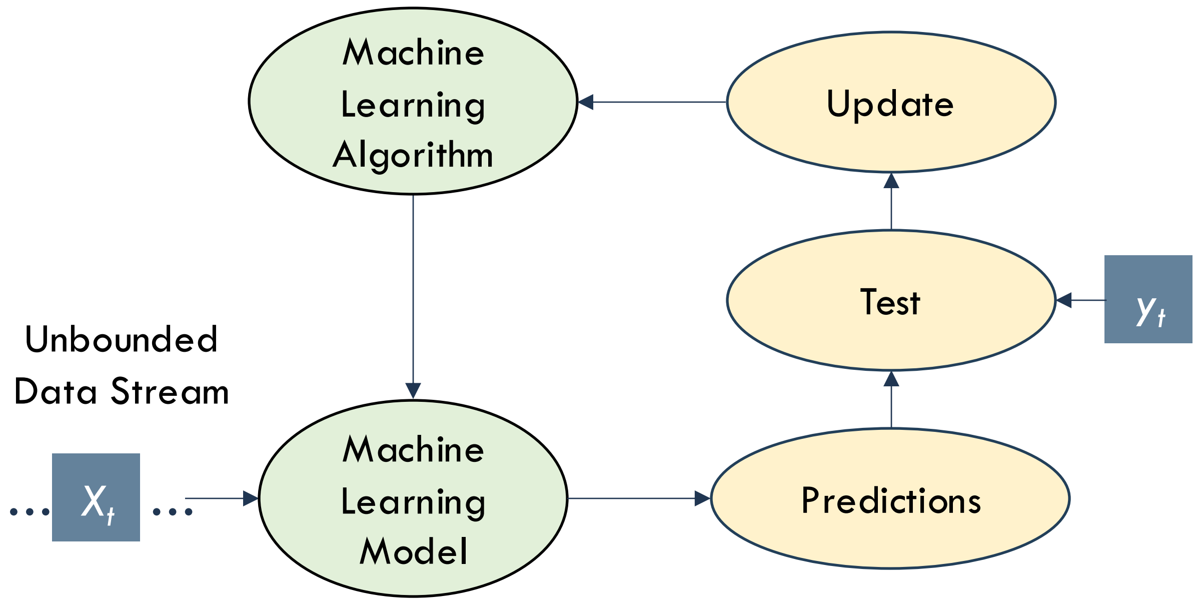
F. Giannini et al.: Addressing Temporal Dependence, Concept Drifts, and Forgetting in Data Streams

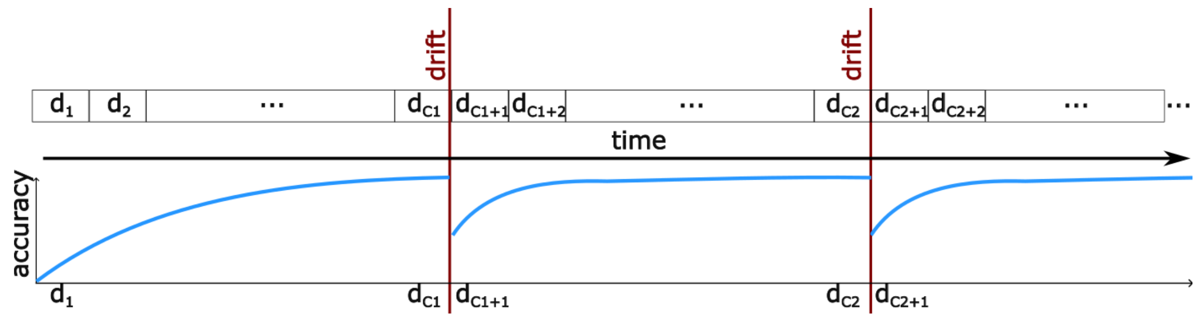
Experimental Setting

- Both synthetic and real weather data streams.
- Two boundary functions alternated (four classification functions).
- Three **abrupt concept drifts.**
- Some drifts are **severe** (more than 50% label changes), and others are **mild.**
- Two studies:
    - cPNN ablation study
    - Comparison between cPNN, SML models (ARF and HAT) and SML models with temporal augmentation.



Evaluation





- Cohen's Kappa Score computed continuously in a prequential evaluation way.
- Score reset after concept drifts.
- Ablation study accumulates data points in mini-batches also for inference.