Online Learning in Changing Environments

Leandro Lei Minku

Centre of Excellence for Research in Computational Intelligence (CERCIA)
School of Computer Science
The University of Birmingham

Outline

- Introduction:
  - Machine learning.
  - Concept drift and changing environments.
  - Online learning.
- Ensemble learning and diversity.
- Research question answered by the work.
- Diversity for dealing with concept drifts.
- Conclusions and future work.

Introduction – Machine Learning

Input features

Output

Introduction – Concept Drift and Changing Environments

- Most learning machines operate in *offline* mode:
  1. Learn a task based on examples.
  2. Are used to perform the task.
- Many real world applications change with time:
  - Information filtering, credit card approval, intrusion detection, spam detection, market basket.
- Concept drift = change.
- Changing environment = can suffer concept drift.

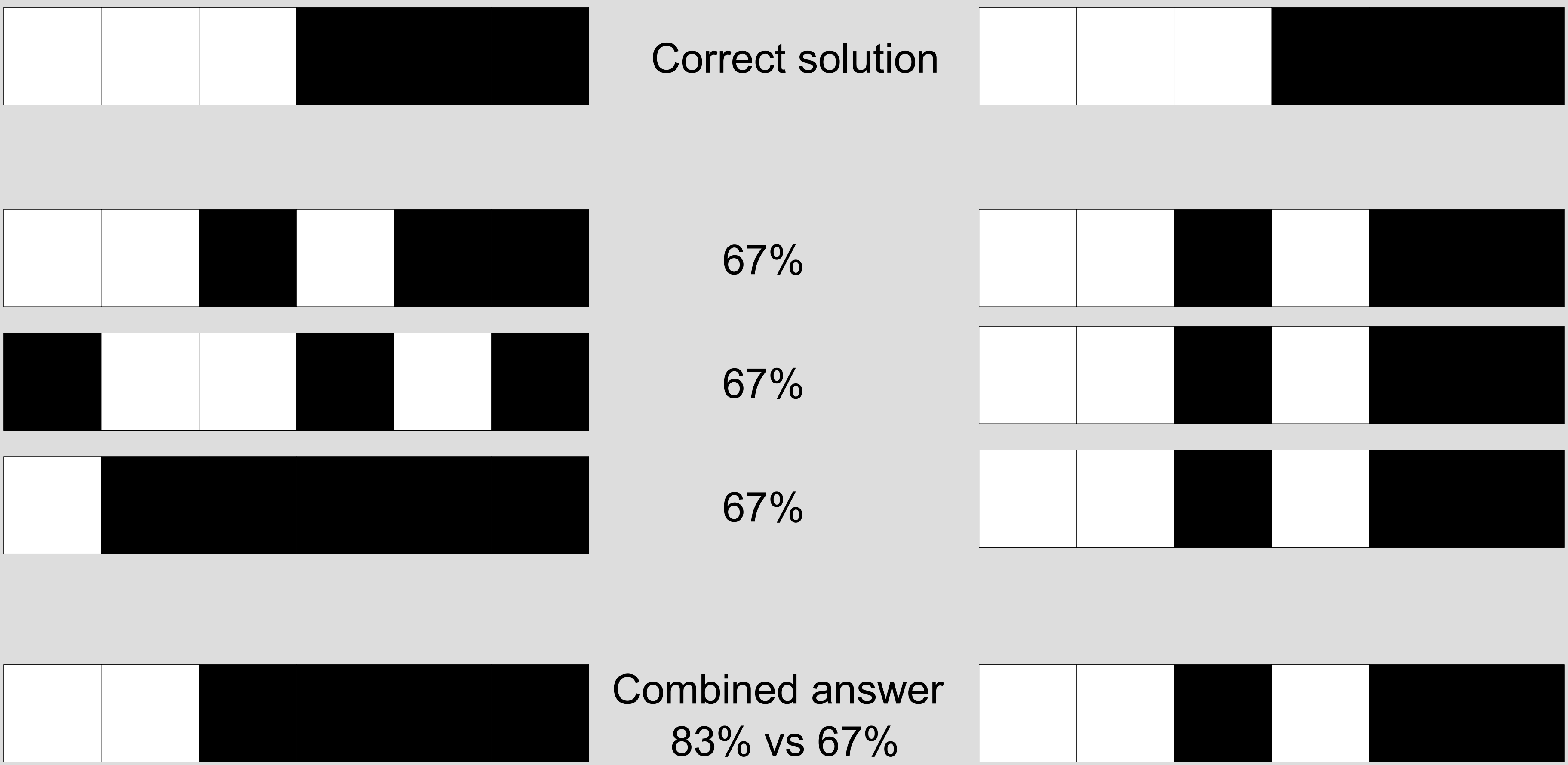
Introduction – Online Learning

- Online learning machines:
  - Learn at the same time as they are being used.
  - Learn during their entire existence.
  - If well designed, they can deal with drifts.
- Online = learn each example separately.
  - Fast, less memory.

Ensembles of Learning Machines

- Ensembles = sets of learning machines.
- Ensembles have empirically and theoretically shown to improve accuracy of single learning machines.
- Diversity has been widely studied in offline learning.

Diversity



Research Question

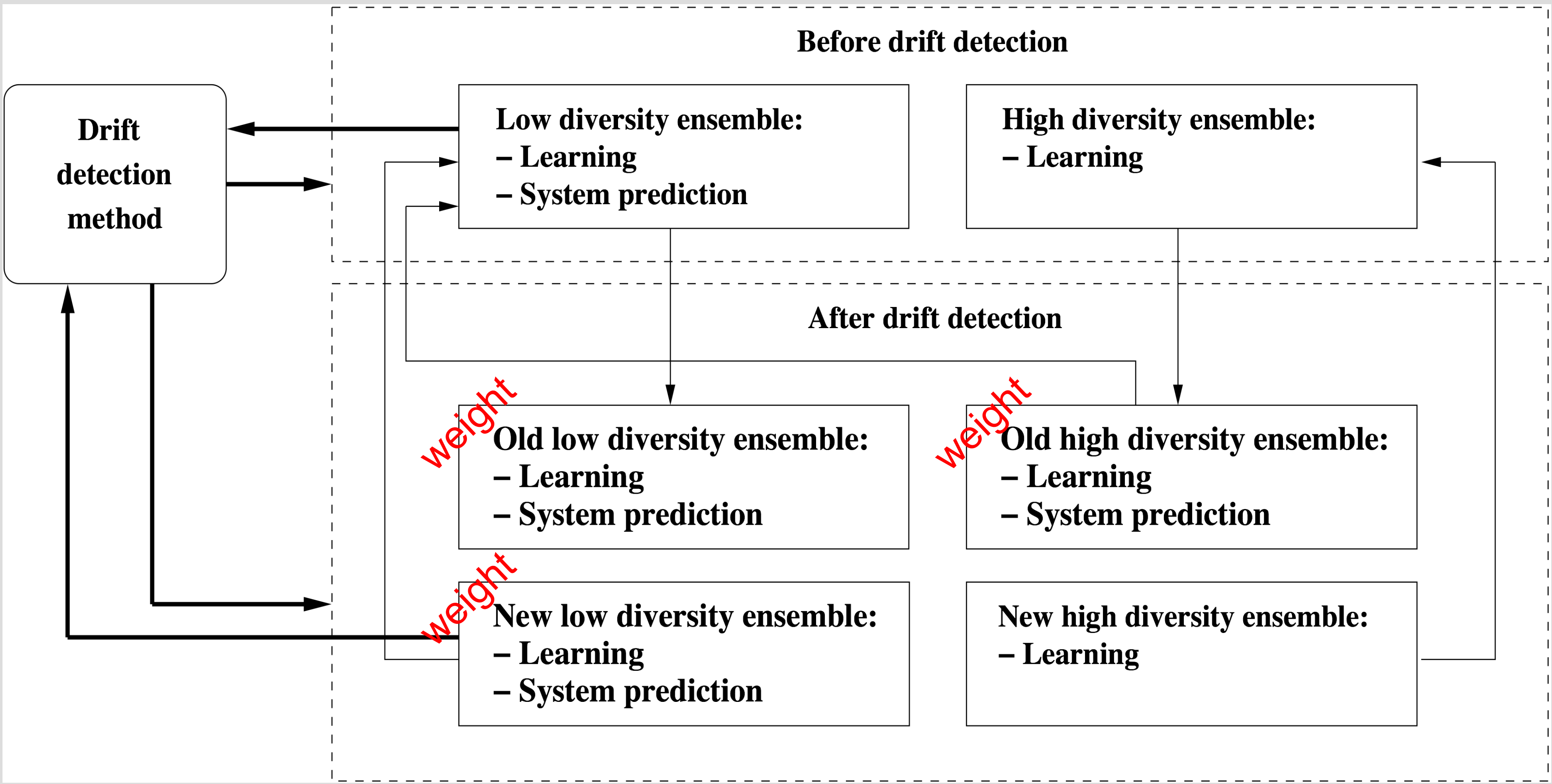
- Ensembles have been used for offline and online learning.
- Ensembles have been used to deal with drifts.
- However, they still suffer when there are drifts, even if the changes are small.

How to use information previously learnt to aid the learning after changes?

Diversity for Dealing with Concept Drifts (DDD)

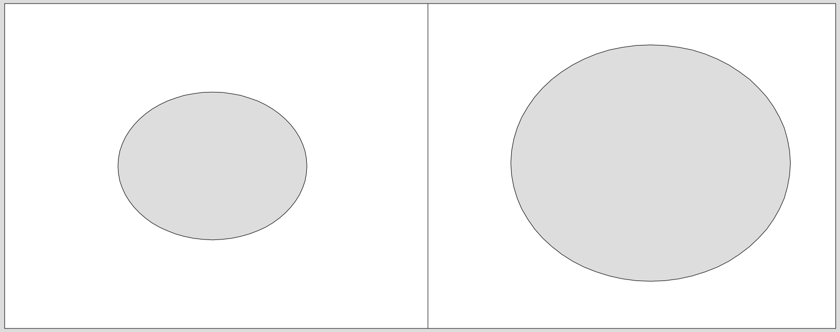
- Problem: learning machines that learnt too well have difficulties to adapt to changes.
- Solution: ensembles with different diversity levels.
  - Allow one ensemble to learn very well by encouraging some diversity.
  - Do not allow the other to learn very well by forcing *a lot* of diversity.
- When there is a drift:
  - Use the high diversity ensemble to adapt to the new situation in hands by using now low diversity.
  - It will be able to use knowledge partially learnt to aid the learning of the new situation.

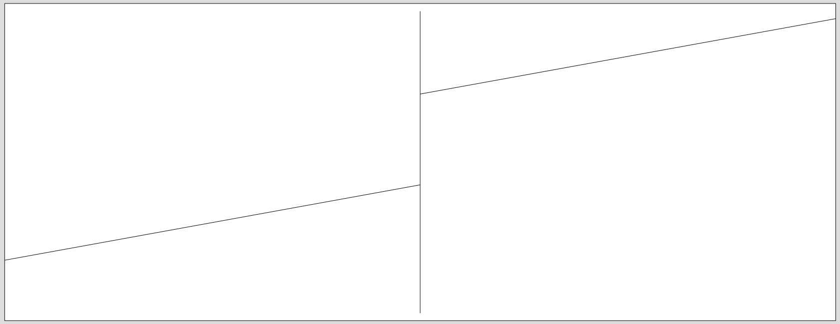
Diversity for Dealing with Concept Drifts (DDD)



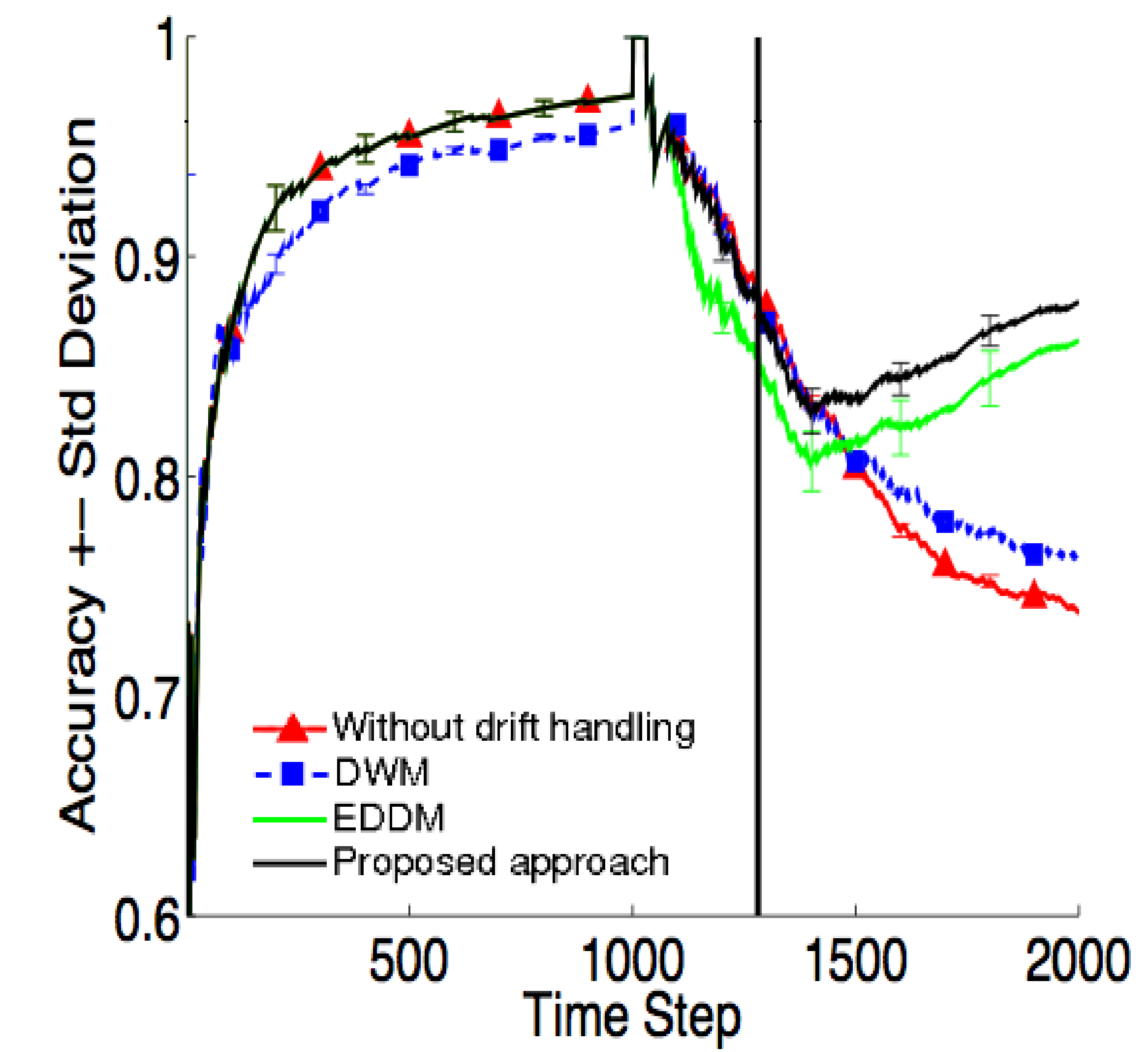
- Online bagging – adapted to present each training example *Poisson(lambda)* times.

Experimental Study – Toy Problems

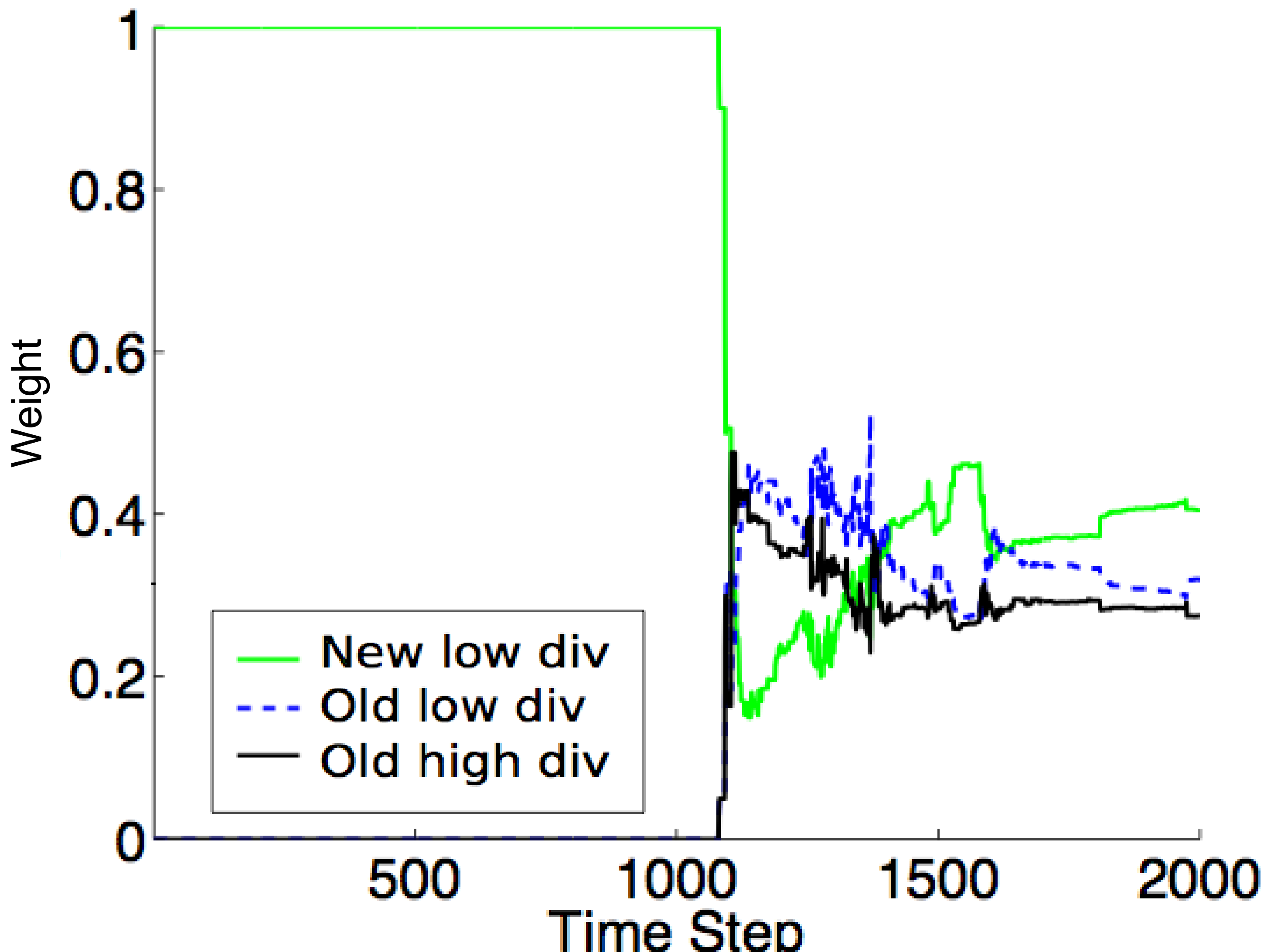




DDD achieves higher accuracy especially for small or slow changes.



Experimental Study – Toy Problems



- Successful selection of old high diversity ensemble.
- Successful use of old low diversity ensemble on false alarms.

Experimental Study

T Tests for Artificial Data – Win / Draw / Loss	
DDD vs EDDM	DDD vs DWM
45% / 46% / 7%	59% / 25% / 15%

- DDD achieves good results due to:
  - The use of information previously learnt to aid learning after drifts.
  - Emphasis on the right ensembles when there are drifts.
  - Robust to false drift detections.
- Experiments with 3 real world problems (credit card, electricity market and intrusion detection) reflect these results.

Conclusions and Future Work

- DDD successfully uses ensembles with different diversity levels:
  - It has at least similar accuracy to other approaches, with very few exceptions.
  - It increases the accuracy in the presence of small or slow changes due to the use of knowledge learnt before changes to *aid* learning after changes.
  - It maintains good accuracy in the absence of drifts by being robust to false drift detections.
- Apply/adapt DDD for specific real world problems.