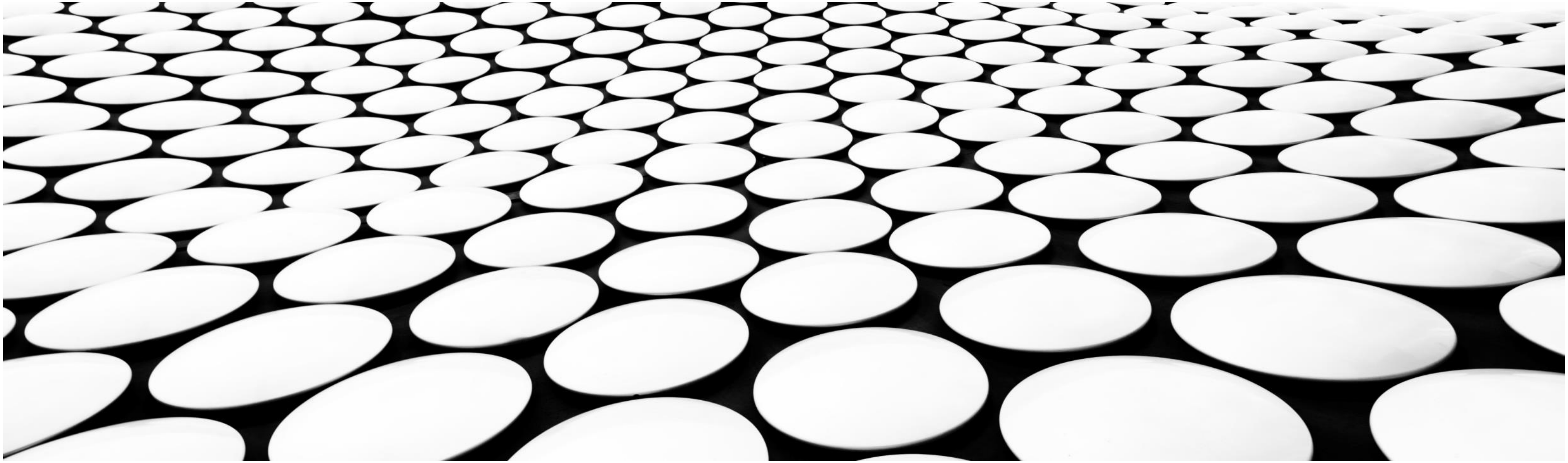
An equivilence Between Private Classification and Online Predicition

Presentation done by : Ayoub Youssoufi, Yassin El hajj chehade



Content Analysis

- What type of paper is it ?

→ Science/ fondamental study

- What is the problem ?

→ Relationship between Online Prediction and Private Classification

- What is the contribution ?

→ Proofs the relation between Online Prediction and Private Classification

- Is it well supported ?

→ The paper is well-supported because he used many references,

→ The paper structure

- What do you like about the paper ?

→ The paper is well structured (overview, main results, proofs, conclusion)

- Is it well written ?

→ It is easy to understand

Private classification

Definition:

Private classification: is a method used for publicly sharing information about a dataset by describing the patterns of groups in the dataset while hiding informations about individuals in the dataset.

An analysis of a dataset is private if the outcome should not have disproportionately large adverse impact on the same dataset.

An example of private classification is DP-PAC learnable model which it has been investigated by demonstrating the equivalence between Online-learnable and Private classification only in binary classification

Online-learning

<u>Definition:</u>

It is well-studied machine learning branch which consists on making prediciton on real-time arriving dataset. This task arise specially on recommandation systems, advertisement placement,…etc

A class H = {h : X ⟶ {±1}} is online learnable if and only if it has a finite Littlestone dimension d where it can determined through the best possible regret.

$$\boxed{\sum_{t=1}^{T} 1[y_t \neq \hat{y}_t] - \min_{h^* \in \mathcal{H}} \sum_{t=1}^{T} 1[y_t \neq h^*(x_t)].}$$

Every finite class H has Littlestone dimension d ≤ log|H|.

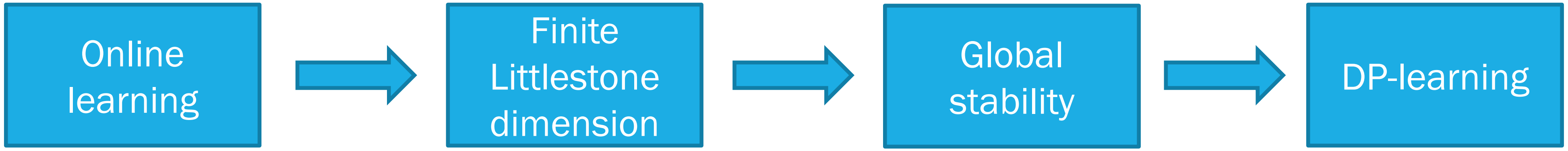
Stability

Online learning and Private classification has from of connection appearing in the global stability.

1. Private classification : it requires the robustness of the distribution of the outcome of an algorithm when its input undergoes small changes.

2. Online learning: is global stable because any classification can be bounded with a Littlestone dimension

Therefore :



Online learning

Finite Littlestone dimension

Global stability

DP-learning

If we have an online learning means it exists a finite « d »Littlestone dimension which ensure the stability of the classification even if the input undergoes small changes.

Global stability & Littlestone dimension

- For η > 0 and n ∈ N, a learning algorithm A is (n, η)-globally stable, with respect to a distribution D over examples if there exists an hypothesis h whose frequency as an output is at least η.

$$\boxed{\Pr_{S\sim\mathcal{D}^n}[\mathcal{A}(S) = h] \geq \eta.}$$

- The argument in the paper shows that every H can be learned by a globally-stable algorithm with parameters η = exp(exp(−d)) and n = exp(exp(d)), where d is the Littlestone dimension of H.

- For the rest of the manuscript, the authors state the main results and discuss some implications. Then provide the complete proofs, and at the end conclude the paper with some suggestions for future work.

Limitations and assumptions

- The paper investigates the case of binary classificaion, without generalizing with the multiclassification or regression
- The investigation was only using the global stabibility, The author propose to explore potential connections to other forms of stability, such as local statistical stability, uniform hypothesis stability, etc,
- The upper bound on the DP complexity of a class H has a double exponential dependence on the Littlestone dimension d. While the lower bound is on log$*$(d). So, for the future we want to improve the upper bound.