

### Efficiently Monitoring Small Data Modification Effect for Large-Scale Learning in Changing Environment

花田 博幸<sup>1</sup>      柴垣 篤志<sup>1</sup>      佐久間 淳<sup>2,3,4</sup>      竹内 一郎<sup>1,4,5</sup>  
Hiroyuki Hanada    Atsushi Shibagaki    Jun Sakuma    Ichiro Takeuchi

出典：The 32nd AAAI Conference on Artificial Intelligence (AAAI 2018), pp. 1314-1321

機械学習において、大規模なデータセットからの学習結果を得たのちに、その一部の値が変化した場合を考える。このとき、変化後のデータセットに対応した学習結果を得るためには、一般にデータセット全体の大きさに比例するという、大きい計算コストが必要となる。本研究では、変化後のデータセットに対する学習結果を厳密に求めることなく、学習結果の変動の上界を高速に、具体的にはデータセットの値が変化した箇所の個数のみに比例する時間で計算することを考える。学習結果の変動を高速に検知できることは、(大きい計算コストが必要な)再学習を行うべき時期を判断するのに利用できるなどの応用がある。

本研究では特に正則化付き経験損失最小化問題を対象とする。そして、正則化関数が強凸である(L2ノルム+凸関数の形でかける)場合や、損失関数が平滑である(勾配がリプシッツ連続である)場合に、学習結果の変動の上界を高速に得る計算方法を示している。具体的な方法としては、解くべき凸最適化問題の双対問題を考え、その“duality gap”を用いて学習結果の変動の上界を与えている。

---

1 名古屋工業大学, Nagoya Institute of Technology

2 筑波大学, University of Tsukuba

3 科学技術振興機構, JST

4 理化学研究所, RIKEN

5 物質・材料研究機構, NIMS