Gradient-based Algorithms for Machine Teaching

Pei Wang
Kabir Nagrecha
Nuno Vasconcelos
SVCL, ECE, University of California, San Diego

Motivation
• Although crowd-sourcing can scalably annotate everyday objects, actions, or scenes data, it is hard to do it on fine-grained expert domain, because annotations require highly specialized and domain specific knowledge.
• Annotation by specialists is usually too expensive and rarely feasible at a large scale.
• Less label-intensive forms of learning, including few-shot learning, transfer learning, semi-supervised learning and self-supervised learning, still underperform supervised learning.
• We use machine teaching algorithms to train crowdsource annotators to label data from specialized domains and make scalable supervised learning possible.

MaxGrad
• Optimal student assumption
  • Mainly focus on crowd sourcing context;
  • The teaching set must be small;
  • Humans are good at few-shot learning scenery;
  • free-willing participants rated by their performance;
• Iterative machine teaching

Preliminaries
• Machine teaching
  • It is assumed that the teacher can access to a much larger example dataset \( \mathcal{D} = \{(x_1, y_1), \ldots, (x_N, y_N)\} \)

Select new examples by MaxGrad
• At iteration \( t \), the teacher has access to the population risk \( \mathcal{R}_\mathcal{D}(f^t) \) and corresponding steepest descent direction;
• The student can only learn from the teaching set \( \mathcal{L}^{t-1} \) of iteration \( t-1 \) and newly selected examples \( \mathcal{N} \);
• MaxGrad selects \( N \) so that the steepest descent direction on \( \mathcal{L} = \mathcal{L}^{t-1} \cup \mathcal{N} \) is closest to \( g^* \).

Experiment results
• On the simulated learners
  • The population is Butterflies and Chinese Characters.
  • The results show that MaxGrad outperforms other methods.
• On the real learners
  • The real learners are turkers who are rated by their performance.

Reference
1. Adish Singla, Ilija Bogunovic, Gabor Bartok, Amin Karbasi, and Andreas Krause. Near-optimally teaching the crowd to classify, ICML 2014
2. Oisin Mac Aodha, Shihan Su, Yuxin Chen, Pietro Perona, and Yisong Yue. Teaching categories to human learners with visual explanations, CVPR 2018
3. Weiyang Liu, Bo Dai, Ahmad Humayun, Charlene Tay, Chen Yu, Linda B Smith, James M Rehg, and Le Song. Iterative machine teaching, ICML 2017
4. Weiyang Liu, Xin Chen, Zhen Liu, James M Rehg, and Le Song. Towards black-box iterative machine teaching, ICML 2018

Algorithm 1 MaxGrad
Input \( \mathcal{D} = \{(x_1, y_1), \ldots, (x_N, y_N)\} \), codewords \( Y \), max iter. \( T \), effort \( \tau \)
1. Initialization: \( \mathcal{L}^0 = \emptyset, f^0, \mathcal{D}^0 = \mathcal{D} \)
2. for \( t = 1, \ldots, T \) do
3.  compute \( \xi_t \) for all examples in \( \mathcal{D}^{t-1} \),
4.  order examples by decreasing \( \xi_t \) and select top \( \tau \) to create \( N_t \),
5.  teaching set update: \( \mathcal{L}^t = \mathcal{L}^{t-1} \cup N_t \)
6.  student update: \( f^{t+1} = f^t(\mathcal{L}^t) \),
7. end for
Output \( \mathcal{L}^T \)