Discussion Leader Cover Letter

JUSTIN DULAY

ACM Reference Format:
Justin Dulay. 2022. Discussion Leader Cover Letter. 1, 1 (February 2022), 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In this discussion of the latest literature, I want to explore the problem of augmenting human intelligence into the training regimes of artificial learning agents. There are many different methods to augment machine learning models with human input data, such as human-in-the-loop objections, biologically inspired algorithms, and psychophysical annotations. In this work, given the nature of my previous research and potential pathways, I will explore relevant work in past exploration of psychophysical annotations into machine learning models — and furthermore, identify research gaps with the hope of making improvements to the field.

2 PAPERS

The three candidate papers that I will list explore the research question of parameterizing ML models and possible pathways to exploit this. Below are the titles and reference links:

- *Perceptual Annotation: Measuring Human Vision to Improve Computer Vision* (https://ieeexplore.ieee.org/abstract/document/6701391)
- *Proximal Policy Optimization Algorithms* (https://arxiv.org/abs/1707.06347)
- I also note other papers that fit within this problem-scope, but outside the scope of discussion for this class (*i.e.* one is just a well-cited tech report).

2.1 Psychophysics

The first paper we examine is through my lab's previous original implementation of parameterizing psychophysical labels into an artificial learning agent training process. It follows the following research question: "We describe methods for incorporating this item response data into the objective function of support vector machines, effectively using human performance to guide and regularize a problem's solution". While not exactly a question, they continue to address the problem of performance gap between humans and machines on perceptual tasks and answer it with perceptual annotation through psychophysics experiments.

The paper claims utilizing data from crowd-sourced versions of traditional psychology psychophysics experiments in the training process of an SVM improves model performance vs without a modified label space. The ML task was a supervised facial recognition task with an SVM. In order to parameterize ML performance on this, they created a psychophysical tasks for crowd-sourced workers on a wide-audience platform. The annotators were asked if there was

Author's address: Justin Dulay.

Manuscript submitted to ACM

a difference between two alternative images, often included with noise to occlude clarity. They stored the reaction time with each item and created a label penalty for a hinge loss function to use with the model.

Their model reported modest improvements on a facial recognition task when compared to the control.

The blending of classical psychological testing as a means of model enhancement in ML is interesting. This sparked interest in utilizing psychology experiments as a means to get a different from of human-in-the-loop performance enhancements for models, without actually being involved.

2.2 Reinforcement Learning

Proximal policy optimization is a model-free reinforcement learning algorithm that removes noisy gradients through a clipped surrogate policy loss function. The paper proposes this new method with some example experiments to demonstrate its superiority in certain benchmark cases.

The loss function is:

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t \left[\min \left(r_t(\theta)\hat{A}_t, \operatorname{clip}\left(r_t(\theta), 1-\epsilon, 1+\epsilon\right)\hat{A}_t \right) \right]$$

where ϵ is a hyperparameter between 0.1 and 0.2 that controls the rate in which advantages and policies change in a surrogate function ahead of the current policy based upon the probability ratio of the incentive r_t.

One interesting facet of the loss function is the ϵ error definer, that smooths the jumping of the loss function in certain cases. It usually acts as an error reduction parameter by removing noise gradients and only collecting some that are over a certain threshold.

The algorithm clips the loss function at a given epoch and allows for an exploration stage a few time steps in advance. By allowing for my exploration than a traditional reinforcement learning paradigm, the algorithm generalizes more readily on complex RL tasks.

The authors compare this algorithm with several others on a MuJoCo robotics simulator for RL performance, with average reward on generalization being the primary metric.

3 CONCLUDING REMARKS

While there are noticeable papers on RL + Psychophysics, we don't see this style parameterization within the policy training stage of reinforcement learning algorithms. In future works, I would like to incorporate the psychophysical parameterization style of the first paper into the learning paradigm described in the second.

I also want to postulate one last point. I don't think either of these papers is worth talking about for a discussion in their own right, but there is a dynamic between PsyPhy — an object perturbation psychophysics framework, published by RichardWebster et al. at Notre Dame — and Psychlab, which is pre-printed from DeepMind around the same time, which has RL agents perform psychophysical tasks in a virtual environment. These two papers have some contention between them, and I would like to create a different version of them at some point.