AutoML-Zero: Evolving Machine Learning Algorithms From Scratch

Esteban Real [*1] Chen Liang [*1] David R. So [1] Quoc V. Le [1]

Abstract

Machine learning research has advanced in multiple aspects, including model structures and learning methods. The effort to automate such research, known as AutoML, has also made significant progress. However, this progress has largely focused on the architecture of neural networks, where it has relied on sophisticated expert-designed layers as building blocks—or similarly restrictive search spaces. Our goal is to show that AutoML can go further: it is possible today to automatically discover complete machine learning algorithms just using basic mathematical operations as building blocks. We demonstrate this by introducing a novel framework that significantly reduces human bias through a generic search space. Despite the vastness of this space, evolutionary search can still discover two-layer neural networks trained by backpropagation. These simple neural networks can then be surpassed by evolving directly on tasks of interest, *e.g.* CIFAR-10 variants, where modern techniques emerge in the top algorithms, such as bilinear interactions, normalized gradients, and weight averaging. Moreover, evolution adapts algorithms to different task types: *e.g.*, dropout-like techniques appear when little data is available. We believe these preliminary successes in discovering machine learning algorithms from scratch indicate a promising new direction for the field.

1. Introduction

In recent years, neural networks have reached remarkable performance on key tasks and seen a fast increase in their popularity [*e.g.* He et al., 2015; Silver et al., 2016; Wu et al., 2016]. This success was only possible due to decades of machine learning (ML) research into many aspects of the field, ranging from learning strategies to new architectures [Rumelhart et al., 1986; LeCun et al., 1995; Hochreiter & Schmidhuber, 1997, among many others]. The length and difficulty of ML research prompted a new field, named *AutoML*, that aims to automate such progress by spending machine compute time instead of human research time (Fahlman & Lebiere, 1990; Hutter et al., 2011; Finn et al., 2017). This endeavor has been fruitful but, so far, modern studies have only employed constrained search spaces heavily reliant on human design. A common example is *architecture search*, which typically constrains the space by only employing sophisticated expert-designed layers as building blocks and by respecting the rules of backpropagation (Zoph & Le, 2016; Real et al., 2017; Tan et al., 2019). Other AutoML studies similarly have found ways to constrain their search spaces to isolated algorithmic aspects, such as the learning rule used during backpropagation (Andrychowicz et al., 2016; Ravi & Larochelle, 2017), the data augmentation (Cubuk et al., 2019a; Park et al., 2019) or the intrinsic curiosity reward in reinforcement learning (Alet et al., 2019); in these works, all other algorithmic aspects remain hand-designed. This approach may save compute time but has two drawbacks. First, human-designed components bias the search results in favor of human-designed algorithms, possibly reducing the innovation potential of AutoML. Innovation is also limited by having fewer options (Elsken et al., 2019b). Indeed, dominant aspects of performance are often left out (Yang et al., 2020). Second, constrained search spaces need to be carefully composed (Zoph et al., 2018; So et al., 2019; Negrinho et al., 2019), thus creating a new burden on researchers and undermining the purported objective of saving their time.

To address this, we propose to automatically search for *whole* ML algorithms using *little* restriction on form and *only* simple mathematical operations as building blocks. We call this approach *AutoML-Zero*, following the spirit of previous work which aims to learn with minimal human participation [*e.g.* Silver et al., 2017]. In other words, AutoML-Zero aims to search a fine-grained space simultaneously for the model, optimization procedure, initialization, and so on, permitting much less human-design and even allowing the discovery of non-neural network algorithms. To demonstrate that this is possible today, we present an initial solution to this challenge that creates algorithms competitive

[*]Equal contribution [1]Google Brain/Google Research, Mountain View, CA, USA. Correspondence to: Esteban Real <ereal@google.com>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020.