

AlphaD3M: Machine Learning Pipeline Synthesis

Iddo Drori

IDRORI@NYU.EDU

Yamuna Krishnamurthy

YAMUNA@NYU.EDU

Remi Rampin

REMI.RAMPIN@NYU.EDU

Raoni de Paula Lourenco

RAONI@NYU.EDU

Jorge Piazzentin Ono

JORGEHPO@NYU.EDU

Kyunghyun Cho

KYUNGHYUN.CHO@NYU.EDU

Claudio Silva

CSILVA@NYU.EDU

Juliana Freire

JULIANA.FREIRE@NYU.EDU

New York University, Tandon School of Engineering and Center for Data Science

Abstract

We introduce AlphaD3M, an automatic machine learning (AutoML) system based on meta reinforcement learning using sequence models with self play. AlphaD3M is based on edit operations performed over machine learning pipeline primitives providing explainability. We compare AlphaD3M with state-of-the-art AutoML systems: Autosklearn, Autostacker, and TPOT, on OpenML datasets. AlphaD3M achieves competitive performance while being an order of magnitude faster, reducing computation time from hours to minutes, and is explainable by design.

Keywords: AutoML, Expert Iteration, Sequence Model, Data Driven Discovery of Models.

1. Introduction

Automatic machine learning (AutoML) aims to learn how to learn. Given a dataset, a well defined task, and performance criteria, the goal is to solve the task with respect to the dataset while optimizing performance. Existing systems have focused on a relatively small set of machine learning primitives, with a few tasks (Feurer et al., 2015), or on a small set of datasets (Chen et al., 2018), or on numerous datasets within specific domains (Olson and Moore, 2016).

DARPA’s Data Driven Discovery of Models (D3M) program pushes this vision further and proposes to develop infrastructure to automate model discovery, i.e., solve any task on any dataset specified by the user. Using a broad set of computational primitives as building blocks, the D3M system should synthesize a pipeline and set the appropriate hyper-parameters to solve a previously unknown data and problem. The D3M system also has a user interface that enables users to interact with and improve the automatically generated results (Blei and Myth, 2017).

Inspired by AlphaZero (Silver et al., 2017), we frame the problem of pipeline synthesis for model discovery as a single-player game (McAleer et al., 2018): the player iteratively builds a pipeline by selecting among a set of actions which are insertion, deletion, replacement of pipeline parts. An inherent advantage of this approach is that at the end of the process, once there is a working pipeline, it is completely explainable, including all the actions and decisions which led to its synthesis. Another advantage is that our approach leverages recent advances in deep reinforcement learning using self play, specifically expert iteration

(Anthony et al., 2017) and AlphaZero (Silver et al., 2017), by using a neural network for predicting pipeline performance and action probabilities, along with a Monte-Carlo Tree Search (MCTS), as illustrated in Figure 1 (left), which takes strong decisions based on the network. The process progresses by self play with iterative self improvement, and is known to be highly efficient at finding a solution to search problems in very high dimensional spaces. We evaluate our approach using the OpenML dataset on the tasks of classification and regression, demonstrating competitive performance and computation times an order of magnitude faster than other AutoML systems.

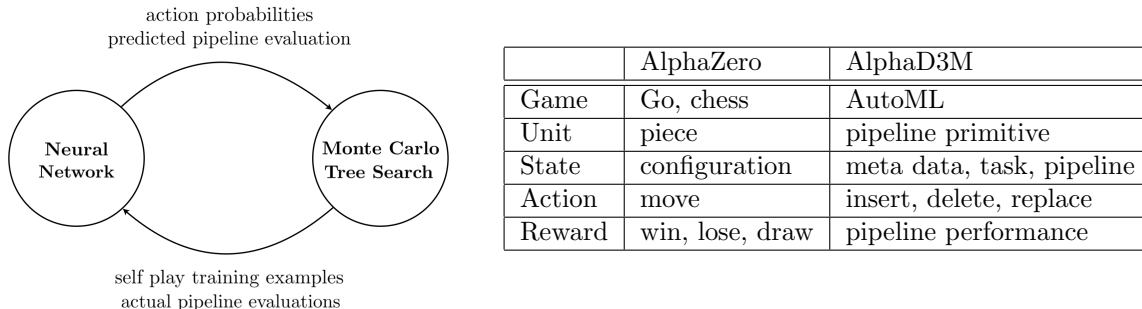


Figure 1: AlphaD3M iterative improvement (left); AlphaD3M game representation (right).

Each of the existing AutoML systems uses any one of the following key elements individually: differentiable programming, tree search, evolutionary algorithms, and Bayesian optimization, to find the best machine learning pipelines for a given task and dataset. Differentiable programming, of which neural network backpropagation is a special case, is used for learning feature extraction and estimation (Ganin and Lempitsky, 2015) and for end-to-end learning of machine learning pipelines with differentiable primitives (Milutinovic et al., 2017). Bayesian optimization methods are used for hyper-parameter tuning (Bergstra and Bengio, 2012). Both AutoWEKA (Kotthoff et al., 2017) and Autosklearn (Feurer et al., 2015) extend the application of these techniques to the selection of the model in addition to the hyper-parameter values, solving the combined algorithm selection and hyper-parameter optimization problem by fitting probabilistic models capturing the relationship between parameter values and performance measures using a Gaussian Process, Random Forest, or tree-structured Parzen estimator (Bergstra et al., 2011). Auto-Tuned Models (Swearingen et al., 2017) represent the search space as a tree with nodes being algorithms or hyperparameters and searches for the best branch using a multi-armed bandit. TPOT (Olson and Moore, 2016) and Autostacker (Chen et al., 2018) uses evolutionary algorithms to generate machine learning pipelines while optimizing their hyperparameters. TPOT represents machine learning pipelines as trees, whereas Autostacker represents them as stacked layers.

Our goal is to search within a large space for the machine learning, and pre and post processing primitives and parameters which together constitute a pipeline for solving a task on a given dataset. The problem is that of high dimensional search. Although the datasets differ, the solution pipelines contain recurring patterns. Just as a data scientist develops intuition and patterns about the pipeline components, we use a neural network along with a Monte-Carlo tree search in an iterative process. This combination results in the network

learning these patterns while the search splits the problem into components and looks ahead for solutions. By self play and evaluations the network improves, incorporating a better intuition. An advantage of this iterative dual process is that it is computationally efficient in high dimensional search (Silver et al., 2017).

2. Methods

Following dual process theory, we solve the meta learning problem by sequence modeling using a deep neural network and Monte Carlo tree search (MCTS) (Silver et al., 2017; Anthony et al., 2017). This section describes our representation, followed by details of the neural network and MCTS.

2.1. Representation

Figure 1 (right) illustrates a high level analogy between a two player competitive game and our single player pipeline synthesis game, including state, action, and reward. A pipeline is a data mining work flow, of pre-processing, feature extraction, feature selection, estimation, and post-processing primitives. Algorithm 2.1 describes our pipeline state representation. Our architecture models meta data and an entire pipeline chain as state rather than individual primitives. A pipeline, together with the meta data and problem definition is analogous to an entire game board configuration. The actions are transitions from one state (pipeline) to another.

Algorithm 1: Pipeline Encoding

Given datasets D , tasks T , and a set of possible pipeline sequences S_1, \dots, S_n , from the available machine learning, and data pre and post processing primitives.

- For each dataset D_i and task T_j :
 1. Encode dataset D_i as meta data features $f(D_i)$.
 2. Encode task T_j .
 3. Encode the current pipeline at time t by a vector S_t .
 4. Encode action $f_a(S_t)$, so policy π maps $(f(D_i), T_j, S_t)$ to $f_a(S_1), \dots, f_a(S_n)$.
-

2.2. Neural Network

AlphaD3M uses a recurrent neural network, specifically an LSTM. Let $f_\theta(s) = (P(s, a), v(s))$, where $P(s, a)$ is the action probabilities and $v(s)$ the evaluation score of the model predicted by the network f with parameters θ , for a given dataset D and task T , for a given state s . The neural network predicts the probabilities over actions a which lead to sequences S that describe a pipeline, which in turn solves the given task on the dataset. The network inputs are training examples (s_t, π_t, e_t) from games of self play, where s_t is the state at time t , π_t the policy estimated by MCTS, and e_t the actual pipeline evaluation at the end of the game. The state s_t is composed of a vector encoded as described in Algorithm 2.1. The network outputs are probabilities over actions $P(s, a)$, and an estimate of pipeline performance v .

We optimize the network parameters θ by making the predicted model S match the real world model R and the predicted evaluation results v match the real world evaluation e , by minimizing the cross entropy loss between S and R , and the mean squared error between v and e . We add an ℓ_2 regularization term for the network parameters θ to avoid over-fitting and an ℓ_1 regularization term which prefers simple pipelines. Thus our network f_θ is trained by minimizing the following non-linear loss function using stochastic gradient descent:

$$L(\theta) = S \log R + (v - e)^2 + \alpha \|\theta\|_2 + \beta \|S\|_1. \quad (1)$$

2.3. Monte Carlo Tree Search

Our algorithm takes the predictions $(P(s, a), v(s))$ of the neural network and uses them in a MTCS by running multiple simulations to search for a pipeline sequence R with a better evaluation. The search result R improves upon the predicted result S given by the network by improving the network policy using the update rule:

$$U(s, a) = Q(s, a) + cP(s, a) \frac{\sqrt{N(s)}}{1 + N(s, a)}, \quad (2)$$

where $Q(s, a)$ is the expected reward for action a from state s , $N(s, a)$ is the number of times action a was taken from state s , $N(s)$ the number of times state s was visited, $P(s, a)$ is the estimate of the neural network for the probability of taking action a from state s , and c is a constant which determines the amount of exploration. At each step of the simulation, we find the action a and state s which maximize $U(s, a)$ and add the new state to the tree if it does not exist with the neural network estimates $(P(s, a), v(s))$ or call the search recursively otherwise. Next, the model represented by R is realized and applied to the data to solve the task, resulting in a better evaluation e which is the result of running the generated pipeline R on the data and task. Thus the real world search provides us with (R, e) , where R is the real world model, consisting of machine learning primitives, and e the real world evaluation of the model and pipeline using those primitives on the data and task.

The neural network predictions, the MCTS model, and the real world evaluation, together, define a loss function shown in Equation 1, which is minimized to improve the neural network parameters. This process continues iteratively until the best model, which automatically solves the task, is found.

Inspired by the neural editor (Guu et al., 2018) we use edit operations that make the pipeline generation explainable by design. For each iteration of self play the MCTS searches the possible valid pipelines. For each state or pipeline the next possible states or pipelines are limited to those derived from the edit operations of the current state.

3. Results

The data consists of 313 different tabular datasets, of which 296 are from OpenML (Vanschoren et al., 2014). We considered classification, both binary (121 datasets) and multi-class (108 datasets), and univariate regression tasks (84 datasets). Baseline pipelines were constructed using sklearn SGD estimators for classification and regression, and an annotated tabular feature extractor which uses linear SVC, Lasso, percentile classification or regression estimators from sklearn.

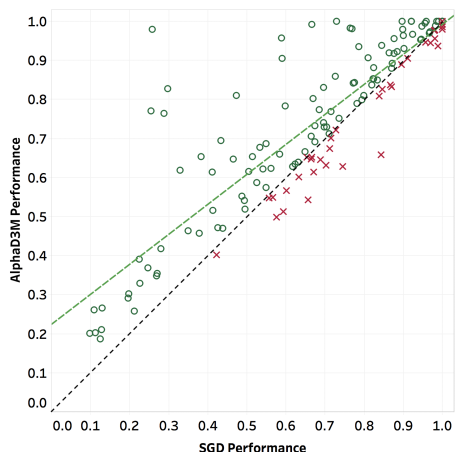


Figure 2: AlphaD3M vs. SGD performance for 180 classification tasks on OpenML datasets.

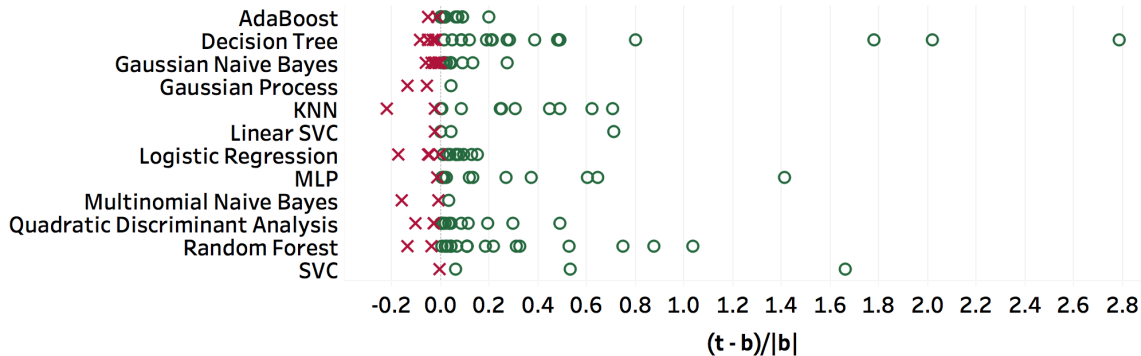


Figure 3: Comparison of normalized AlphaD3M performance t with SGD baseline performance b by estimator.

Figure 2 compares performance between AlphaD3M and SGD which is the baseline pipeline. Each of the 180 points represents a classification task on a different OpenML dataset. The datasets for which AlphaD3M performs better than SGD are shown by green circles and those for which SGD performs better are shown by red crosses. Figure 2 shows that AlphaD3M performs better than baseline for 75% of the datasets, both are comparable for 18% of the datasets, and performs worse for only 7% of the datasets. Figure 3 shows the normalized difference in cross validation performance of AlphaD3M t and SGD baseline b for a classification task for 180 datasets, split according to the estimators used by AlphaD3M, demonstrating better performance across diverse estimators.

Figure 4 compares performance between different AutoML methods: Autosklearn, TPOT, and Autostacker, and our method AlphaD3M, for a number of common OpenML datasets, which serve as representative benchmark datasets for AutoML systems (Olson et al., 2017; Olson and Moore, 2016; Chen et al., 2018). For each method and dataset, we compute the performance mean and standard deviation by repeated evaluation. As shown in Figure 4 our method, AlphaD3M, is competitive with other approaches. All four methods are com-

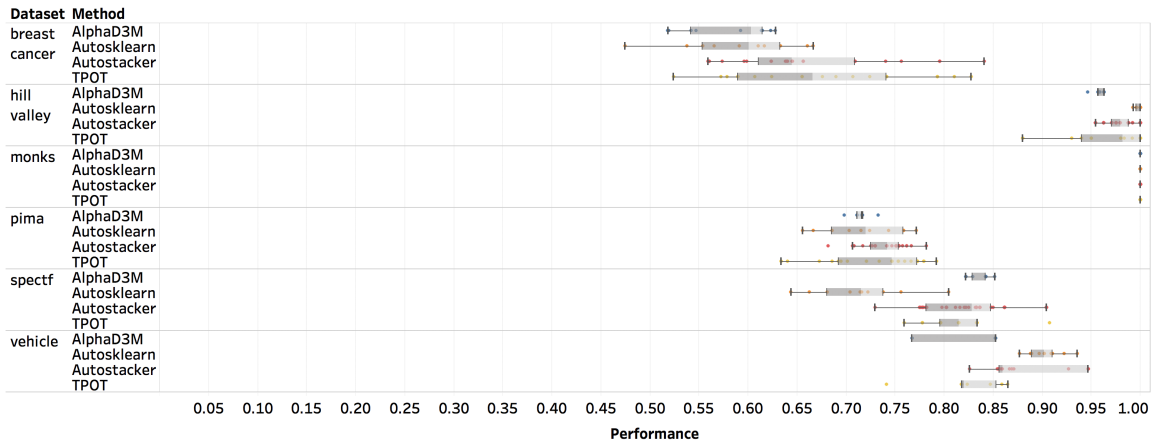


Figure 4: Comparing between performance of AutoML methods on OpenML datasets.

Table 1: Running time comparison (in seconds and speedup factors).

Dataset/Method	TPOT	Autostacker	AlphaD3M	Speedup vs TPOT	Speedup vs AS
breast cancer	3366	1883	460	7.3	4
hill valley	17951	8411	556	32.2	15.1
monks	1517	1532	348	4.3	4.3
pima	5305	1940	619	8.5	3.1
spectf	4191	1673	522	8	3.2
vehicle	16795	4010	531	31.6	7.5

petitive and on par, as their performance including confidence intervals intersect; whereas SGD and Random Forest are not competitive with the leading AutoML methods. We make results for these and other datasets and methods available online ([Drori et al.](#)).

AlphaD3M is implemented using PyTorch. Our implementation takes advantage of GPUs while training the neural network and uses CPUs for the MCTS. Table 1 compares the running time of TPOT, Autostacker, and AlphaD3M on the same datasets, along with the corresponding speedup factors. Table 1 shows that AlphaD3M performs on average an order of magnitude faster, reducing computation time from hours to minutes.

4. Conclusions

We introduced AlphaD3M, an automatic machine learning system with competitive performance, which is an order of magnitude faster than existing state-of-the-art AutoML methods, reducing computation time from hours to minutes. We presented the first single player AlphaZero game representation applied to meta learning by modeling meta-data, task, and entire pipelines as state.

Acknowledgements

This work has been supported in part by the Defense Advanced Research Projects Agency (DARPA) Data-Driven Discovery of Models (D3M) Program.

References

- Thomas Anthony, Zheng Tian, and David Barber. Thinking fast and slow with deep learning and tree search. In *Conference on Neural Information Processing Systems*, 2017.
- James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 2012.
- James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. *Conference on Neural Information Processing Systems*, 2011.
- David M. Blei and Padhraic Myth. Science and data science. In *Proceedings of National Academy of Sciences*, 2017.
- Boyuan Chen, Harvey Wu, Warren Mo, Ishanu Chattopadhyay, and Hod Lipson. Autostacker: A compositional evolutionary learning system. *The Genetic and Evolutionary Computation Conference*, 2018.
- Iddo Drori, Yamuna Krishnamurthy, Remi Rampin, Raoni de Paula Lourenco, Jorge Pizantin Ono, Kyunghyun Cho, Claudio Silva, and Juliana Freire. AlphaD3M results repository. URL <https://osf.io/4hfzm/>.
- Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Tobias Springenberg, Manuel Blum, and Frank Hutter. Efficient and robust automated machine learning. In *Conference on Neural Information Processing Systems*, 2015.
- Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. *International Conference on Machine Learning*, 2015.
- Kelvin Guu, Tatsunori B Hashimoto, Yonatan Oren, and Percy Liang. Generating sentences by editing prototypes. *Transactions of the Association for Computational Linguistics*, 2018.
- Lars Kotthoff, Chris Thornton, Holger H Hoos, Frank Hutter, and Kevin Leyton-Brown. Auto-WEKA 2.0: Automatic model selection and hyperparameter optimization in WEKA. *The Journal of Machine Learning Research*, 18(1), 2017.
- Stephen McAleer, Forest Agostinelli, Alexander Shmakov, and Pierre Baldi. Solving the Rubik’s cube without human knowledge. *arXiv preprint arXiv:1805.07470*, 2018.
- Mitar Milutinovic, Atlm Gunes Baydi, Robert Zinkov, William Harvey, Dawn Song, and Frank Wood. End-to-end training of differentiable pipelines across machine learning frameworks. In *International Conference on Learning Representations*, 2017.
- Randal Olson and Jason Moore. TPOT: A tree-based pipeline optimization tool for automating machine learning. In *International Conference on Machine Learning*, 2016.
- Randal S Olson, William La Cava, Patryk Orzechowski, Ryan J Urbanowicz, and Jason H Moore. Pmlb: a large benchmark suite for machine learning evaluation and comparison. *BioData mining*, 2017.

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *Conference on Neural Information Processing Systems*, 2017.

Thomas Swearingen, Will Drevo, Bennett Cyphers, Alfredo Cuesta-Infante, Arun Ross, and Kalyan Veeramachaneni. ATM: A distributed, collaborative, scalable system for automated machine learning. In *IEEE International Conference on Big Data*, 2017.

Joaquin Vanschoren, Jan N Van Rijn, Bernd Bischl, and Luis Torgo. OpenML: networked science in machine learning. *ACM SIGKDD Explorations Newsletter*, 15(2), 2014.