10.18686/aitr.v2i3.4419

AutoML Algorithm Framework Based on Reinforcement Learning

Mingwei Xu

Zhejiang University of Technology, Hangzhou, Zhejiang 310000

Abstract: Aiming at the existing problems in machine learning, a new machine learning method based on reinforcement learning is proposed. This paper designs and implements the machine learning algorithm in detail from three stages: meta-learning, reinforcement learning and ensemble learning.

Keywords: Reinforcement learning; Algorithm framework; Learning stage

Introduction

Existing machine learning algorithms and systems have proposed different solutions for automating machine learning processes. Auto-Weka and Auto-Sklearn et al. adopted the Bayesian optimization algorithm SMAC based on random forests. However, when the number of candidate samples increases, the resources required for its initialization will also increase, resulting in a decrease in optimization efficiency; and the iteration time of the Bayesian algorithm will continue to increase with the increase of the parameter space. TPOT regards the machine learning process as a tree pipeline and optimizes it using a genetic algorithm, but its characteristic is that it requires repeated iterations and lacks a reliable convergence theory basis [1]. To solve the above problems, this project intends to study machine learning methods based on reinforcement learning to realize the automated design of machine learning processes. On this basis, a new machine learning method based on Q-Leaming is proposed. This paper adopts a method that combines meta-learning and ensemble learning to optimize the automatic generation of machine learning pipelines.

1. Overall Framework

Figure 1 shows the overall architecture of the reinforcement learning proposed in this paper, which includes a three-stage automatic machine learning algorithm. The method consists of three parts: meta-learning, reinforcement learning, and ensemble learning. The automatic generation algorithm of machine learning pipeline requires the user to provide a standardized data set and information related to the data set, such as the column where the classification features are located, the size of the data, etc. [2]. On this basis, through the three-stage learning process of meta-learning, reinforcement learning, and ensemble learning, the mapping from raw data to the final model is achieved, providing a way for users to understand machine learning algorithms from a black box perspective.

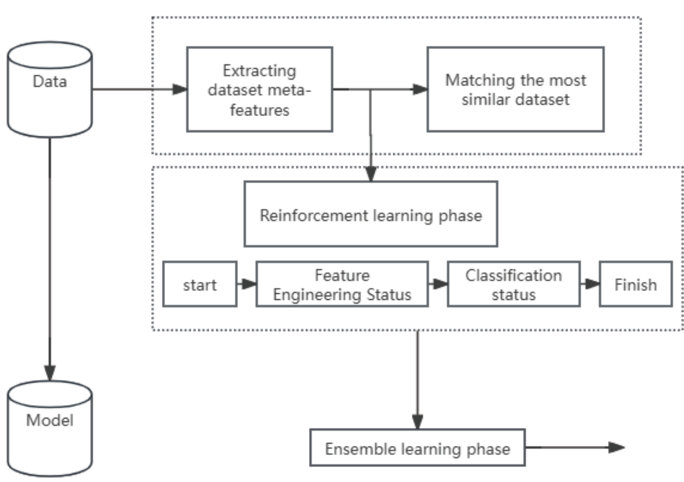

Figure 1 Overall framework of the three-stage automated machine learning algorithm based on reinforcement learning

The specific contents include: meta-learning to extract meta-features from the data set, matching the closest samples on this basis, and extracting the running information (Q-Table) on the samples, replacing the original Q-table with a customized Q-table to speed up the convergence of the algorithm.

The task of the reinforcement learning stage is to automatically search for the machine learning pipeline with the best performance. The state space is constructed using the candidate machine learning algorithm, and the topological relationship between the state spaces is established on this basis. The Q-learning algorithm is used to study the optimization pipeline [3]. This paper studies the decision-weight-based adaptive learning algorithm to meet the automatic learning tasks under different scales of data in order to solve the time constraint problem in automatic machine learning.

Ensemble learning refers to improving the performance and robustness of the system by using ensemble learning based on multiple completed machine learning models. In the process of reinforcement learning, multiple trained models will be saved. If only a single model with the best prediction effect on confirmed data is selected, a large amount of other models will be wasted. Combining multiple models for prediction can better improve the final prediction effect.

2. Learning Phase

2.1 Meta-learning stage

The core idea of the meta-learning module is that the same data usually comes from the same field, and the machine learning model that performs well in similar data sets will also have higher performance. The meta-feature inference of the sample is performed through the meta-learning method, the closest sample is selected from the offline database, and the running status information (Q-Table) of the sample is extracted, which is used as the basis for the initial Q-table in the enhancement stage. This study is based on the Q-Leaming algorithm, and the Q-table obtained through long-term training includes the running information of the sample. In the meta-learning stage, the closest Q-table is used instead of the random Q-table to improve the convergence speed of the algorithm. Figure 2 shows the whole process.

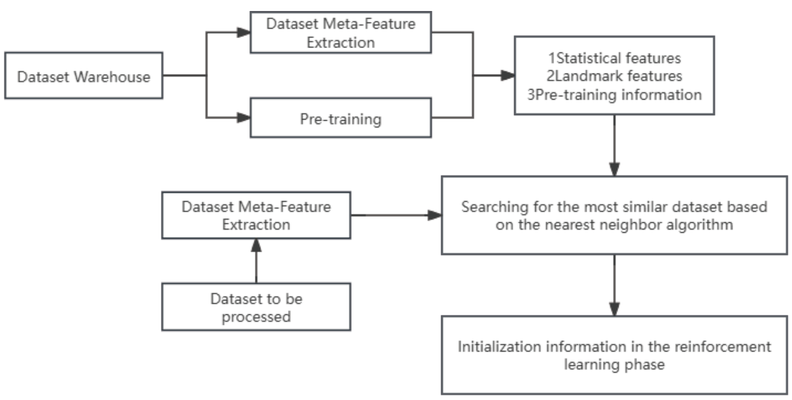

Figure 2 Overall process of meta-learning

The data used in traditional machine learning methods are usually multi-dimensional matrices, with different sample sizes, different feature types, etc. Each data set has its own characteristics. By extracting its meta-features, we can obtain its meta-features, namely, statistical characteristics and landmark characteristics. The so-called statistical characteristics refer to the number of samples, the number of features, the number of features, etc. At the same time, we use machine learning methods to obtain landmark features. For example, the prediction ability of decision trees for data sets can be used as landmark features [4].

The goal of reinforcement learning is achieved through three stages of reinforcement learning. That is, based on the Q-Table in the reinforcement learning process, the Q-Table is continuously updated according to the feedback signals of different mode samples; the values in the Q-Table reflect the performance of different modes in this data set, so we will retain the key operational information of the Q-Table. Finally, by processing the features of the existing data set and the table, metadata is generated and stored in the metadata database.

2.2 Reinforcement Learning Phase

2.2.1 Automatically build machine learning pipeline algorithms

This article divides the machine learning process into five states: initial state, data preprocessing state, feature engineering state, classification state and final state. Figure 3 describes the transformation relationship between each state.

In the re-entry learning phase, the Q-values in the Q-Table are used to characterize the advantages and disadvantages of each state in a specific operation. Each time an operation is executed (adding certain algorithms to the pipeline), the corresponding Q-value is updated. Be-

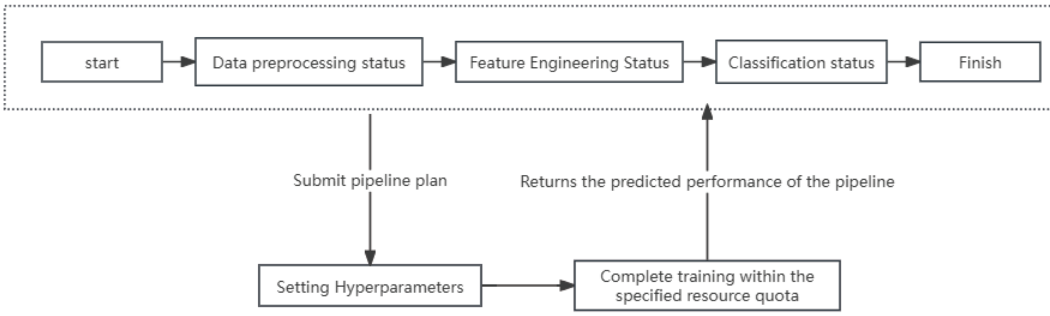

Figure 3 Conversion relationship between machine learning pipeline states

cause the total length of the machine learning pipeline is 102-104, we use a table method to store all Q values. This table that stores Q values is called a Q-table. On this basis, we propose a method based on a machine learning pipeline that uses the performance of the established pipeline in the test set and updates the Q in the Q-table.

2.2.2 Action selection strategy to adapt to time quota

Traditional Q-Leaming uses the ε-greedy algorithm for behavior selection. In the ε-greedy algorithm, the individual will choose a random behavior with probability $, or greedily choose the behavior with the maximum Q in the current state. The probability ε is an important indicator for measuring behavioral decisions and decision decisions. The lower the ε value, the more likely the decision behavior is to make greedy decisions. However, the biggest limitation of automatic machine learning is the limited time resources. A simple idea is that with enough time, people can obtain more information in a variety of ways, and in a limited time, people will try to choose the behavior that can maximize the benefits, and will not waste resources due to random behavior. The traditional ε-greedy method is not well suited to delay-constrained situations because ε is fixed [5].

In the process of automatic learning, due to insufficient information, the system cannot accurately understand the performance of each model on the current data, so it is necessary to explore the performance of various models as much as possible, which requires the search rate to be as high as possible. However, as the number of models increases, the information obtained becomes richer, and the system will also make a rough evaluation of the performance of each model. At this point, you should strive to develop better pipelines while minimizing the use of ordinary pipelines. This method can gradually reduce the search speed as the time limit continues to decrease until it reaches a certain fixed value, thereby effectively solving the time resource limitation problem faced in automatic learning tasks.

2.3 Ensemble learning stage

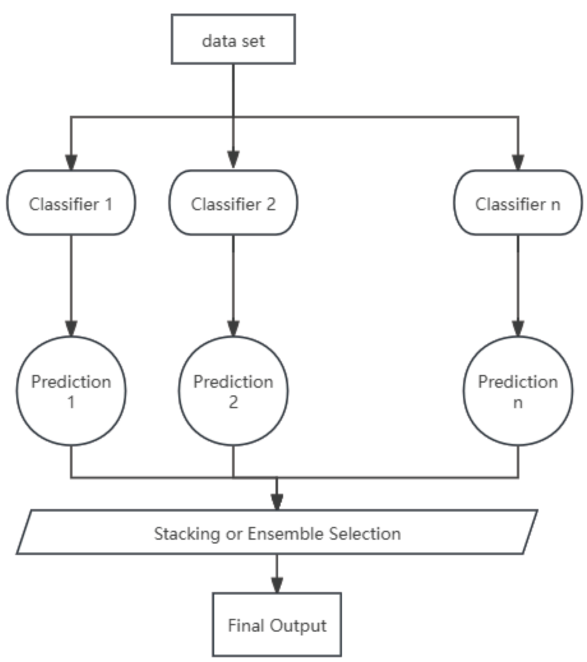

Figure 4 Workflow of the ensemble learning phase

In the reinforcement learning stage, the trained model is stored on the hard disk to improve the final prediction effect. The most direct method is to use the single sample with the best performance among the tested samples as the final sample. This method has problems such as "overfitting" and "waste". In response to the above problems, this article conducts the following research: In the integrated learning stage, two integration methods, Stacking (Stacking) and Ensemble (Ensemble) Selection (Selection), are used to achieve the integration of multiple models. Figure 4 shows the integrated learning process provided in this article.

Among them, the choice of meta-learning algorithm is a very critical issue. In real life, the commonly used method is to use logistic regression as a meta-learning tool, but this method cannot meet the requirements of automatic machine learning.

The Stacking algorithm usually has good prediction results, but it also has the problem of "overfitting". In addition, the Stacking algorithm needs to generate secondary data through interactive confirmation. Therefore, during the training and prediction process of the initial learning algorithm, its operation speed will become very slow.

3. Conclusion

The overall architecture of three-stage learning is proposed, and the important roles of meta-learning, reinforcement learning and ensemble learning in three-stage learning are discussed. Then the whole process of meta-learning is described. On this basis, an idea based on reinforcement learning is proposed to transform the automatic construction problem in the machine learning pipeline into a reinforcement learning problem, and the Q-Leaming technology is used to realize the automatic generation of the model.

References

[1] Huang Lei. Using AutoML technology to promote artificial intelligence education in K-12[J]. Journal of Guangxi Radio and Television University, 2022, 33(01): 24-31.

[2] Zhang Quan, Zhao Guoyi, Su Yuan, et al. Electric power text semantic recognition algorithm based on improved Bert-AutoML [J]. Electronic Design Engineering, 2024, 32(04): 43-46+51.

[3] Zhou Tianbao. Research on automatic machine learning search framework [J]. Information Technology and Standardization, 2021, (07): 33-37.

[4] Zhang Xuefeng, Lin Xiaofei. Automatic machine learning algorithm based on rough set theory [J]. Journal of Shenyang University (Natural Science Edition), 2022, 34(06): 468-477.

[5] Zhang Huajin, Wu Shunchuan, Zhang Zhongxin, et al. Research on automatic machine learning prediction method of slope stability [J]. China Safety Production Science and Technology, 2023, 19(01): 35-40.