Machine learning in economic planning: ensembles of algorithms

J An¹, A Y Mikhaylov² and N E Sokolinskaya²

¹College of Business, Hankuk University of Foreign Studies, 107, Imun-ro, Dongdaemun-gu, Seoul, 130-791, Korea
² Research Center of Monetary Relations, Financial University under the Government of the Russian Federation, 49, Leningradsky Avenue, Moscow, 125468, Russia
³Department of financial markets and banks, Financial University under the Government of the Russian Federation, 49, Leningradsky Avenue, Moscow, 125468, Russia

E-mail: ayumihajlov@fa.ru

Abstract. The algorithm for machine learning of a transport type model is presented for the optimal distribution of tasks in safety critical systems operating in an automatic mode without operator participation. Safety critical systems in various application areas can operate in a wide range of modes - from pure manipulation by the operator prior to their autonomous execution of tasks as part of heterogeneous group. As it is shown by simulation studies of the adaptation algorithm generalized payment matrix of the transport model to the real preferences of the decision maker, even in conditions of significant noise measurements, the proposed algorithm for machine learning model leads to a fairly rapid convergence of estimates. Normalized error from the 15th step does not exceed 10 percent. In this case, the rate of convergence of estimates is not an end in itself in the case of adaptive distribution of tasks in the group of algorithms; an important indicator is the convergence of solutions that exist above the convergence of estimates.

1. Introduction

Machine learning algorithms are used in several fields besides computer science, including critical systems.

Learning is the process of acquiring knowledge. On the contrary, computers do not learn to reason, but improve their algorithms. Nowadays, a large number of algorithms have been proposed in the literature.

Supervised learning occurs when algorithms are equipped with training data and correct answers. The task of the algorithm is to learn from training data and apply the knowledge that was obtained from real data.

In the modern literature the following machine learning algorithms for solving problems in critical systems are presented:

- Bayes Approach
- Decision tree
- Matrix factorization
- Neighborhood
- Neural network
2. Literature review

For ensembles to be more accurate than any of their individual members, the basic models must be as diverse as possible. In other words, the more information comes from the basic classifications, the higher the accuracy of the ensemble.

Some ensemble methods do not define an integrator. But for those methods that do, there are three types of unifiers.

Untrained (Nontrainable). An example of such a method is a simple "major vote" ("major voting").

Trained (Trainable). This group includes "weighted majority voting" and "Naive bays", as well as the "classifier selection" approach, in which the decision on this object is made by one classifier of the ensemble.

Meta-classifier (Meta classifier). The outputs of the base classifiers are considered as inputs to the new trainable classifier, which becomes the unifier. This approach is called "complex generalization," "generalization through learning," or simply stacking. Building a training set for a meta-classifier is one of the main problems of this unifier.

In parallel (independently) or in series you need to train the base classifiers? An example of sequential learning is AdaBoost, where the training set of each added classifier depends on the ensemble created before it.

The following options are available.
1. To manipulate the parameters of learning. Use different approaches and parameters when training individual basic classifiers. For example, it can trigger the neuron weights of the hidden layers of the neural network of each base classifier of the various random variables.
2. To manipulate samples — take your bootstrap sample from the training set for each member of the ensemble.
3. To manipulate the predictors for each base classifier to prepare a different set of predictors determined by a random way. This is the so-called vertical split of the training set.

Is the ensemble constructed by simultaneous training of the required number of classifiers or iteratively, by adding/removing classifiers? Possible options are as follows
1. Quantity reserved in advance
2. The number is set during the training

There is an overproduction of classifiers and their subsequent selection.

Some ensemble approaches can be used with any model of classifiers, while others are bound to a certain type of classifier. An example of a "classifier-specific" ensemble is a random forest (Random Forest). Its base classifier is the decision tree. So, two options approach:
1. Only a specific model of the base classifier can be used;
2. Any model of the base classifier can be used.

When training and optimizing the parameters of the classifier ensemble, it is necessary to distinguish between solution optimization and coverage optimization.

The optimization of decision-making references to the selection of the combiner for a fixed ensemble of base classifiers.
Alternative coverage optimization references to the creation of a variety of base classifications with a fixed unifier.

This decomposition of the ensemble structure reduces the complexity of the problem, so it seems reasonable.

To solve the tasks means to find the set of values of the elements matrix the number of resources (in the classical tasks, resources are homogeneous goods) moved from points of departure to destinations.

The matrix of variables is called the translation plan. An assignment table, in which variables can only take boolean values and in the context of assigning tasks in the group of algorithms means assignment of the tasks to the j-algorithm [1, 2].

The classical criterion for the optimal plan is a minimal total costs. As the initial data usually it is necessary famous vector \( a = [1,2,3...]\) inventory in starting point of tasks, where \( T \) is transpose symbol; vector \( b = [1,2,3...] \) is requirements for each destination.

Matrix C of the cost of transporting a unit of goods. For MASCS integral the costs of the execution of the task are m algorithm [3, 4].

Traditionally tasks decided by criterion of minimization of total costs. In this case, the obtained optimal solution, naturally, will be optimal up to a criterion of the optimality [5].

And since the reality is multi-criteria for almost all applications, the real effectiveness of the solution may be far from the level that satisfies the decision maker as a subject goal setting, or a person who knows the desired level of MASCS effect in any current point in time [6, 7].

This is where the main contradiction arises. Constraining effective use of the task model (and other similar associated) with the multicriteria need of applications and single-digit the material possibility of traditional models [8, 9].

One of the ways to overcome this contradiction is the replacement of the normative scheme of building a model, the essence of which is in identifying some generalized scalar of objective function that approximates the vector of objective functions. The decision maker (explicit and implicit) would transmit it to MASCS for execution [10–12].

Then the objective function plays the role of a formalized image of personal criterion preferences of decision makers.

3. Results and discussion

The authors give a mathematical formulation of the transport model in the normative form, and then they will show the features of building its adaptive (machine learning) option.

One of the varieties of linear programming problems has historically been divided into an independent group due to its special digital structure that allows more effective solution of it with the help specially designed methods for manual calculation.

However, modern software and computing tools allow using standard means of solving.

Next, the authors show how to present the original statement of tasks in the form of standard task. This circumstance allows calling for an adaptive version of the task, based on similar means of linear machine learning [1].

1. Situations represented as a combination of two vectors, and decision-making, which, based on current values payment matrix, resulting in matrix C. Such a task will be called direct task.

2. The found solution is realized, as a result of which the decision making is effect L.

3. The decision-maker (or the person authorized to evaluate decisions) gives an estimate of the accepted solutions \( q \in \{0; 1\} \) is good or bad (i.e. optimal or non-optimal in his opinion).

4. According to the totality of data, adjusted payment matrix values that become current for the following planning step (assignment of tasks).

It remains only to perform distribution of tasks by solving tasks.

We present the formulation of direct and inverse tasks [13, 14] in the form.
\[ R = \left( \sum_{i=1}^{m} \sum_{j=1}^{n} \left( \sum_{t=1}^{K} B^t e_{ij} \right)^2 \right)^{-1} \sum_{t=1}^{K} B^t e_{ij} \] (1)

Where \( e \) are the coordinates of the normal vector of unit length, which are scaled (reduced to a unit length) the coordinates of the vector (matrix) of estimates, \( B \) are weights, from of the next observation calculated by as the length of the observation vector before its normalization. Model example consist of a team of three types of algorithms (\( n = 3 \)). Every algorithm should move to one of two given target points (\( m = 2 \)). Let be simulated several (in the performed simulation experiment - 50) cyclograms of work performance. In each sequence number of algorithms of one type or another can vary from 1 to 7, and the total number of algorithms to perform work at a specific target point may vary in the same range.

The balance of the required and available quantity algorithms complied with transport task is [15, 16].

Decision maker for each cyclogram solves the distribution problem, based on its experience.

4. Conclusion

The research introduced a new algorithm that uses special data transformation for critical system problems.

Safety critical systems in various application areas can operate in a wide range of modes - from pure manipulation by the operator prior to their autonomous execution of tasks as part of heterogeneous group. As it is shown by simulation studies of the adaptation algorithm generalized payment matrix of the transport model to the real preferences of the decision maker, even in conditions of significant noise measurements, the proposed algorithm for machine learning model leads to a fairly rapid convergence of estimates [17, 18].

Normalized error from the 15th step does not exceed 10 percent. In this case, the rate of convergence of estimates is not an end in itself in the case of adaptive distribution of tasks in the group of algorithms; an important indicator is the convergence of solutions that exist above the convergence of estimates.

The algorithm is suitable for estimating tolerance intervals as well as predicting the probability of exceeding arbitrary predefined thresholds. The authors rigorously derived formulas for the probability of crossing the upper bound and the lower bound connected separately (one-sided intervals) or together (two-interval). The experimental evaluation confirmed satisfactory implementation of the proposed technique in terms of the accuracy of the model and the success rate at sufficiently large training sets [19].

The proposed approach can be used in various fields to predict the probability of dangerous situations caused by important system variables that exceed predefined security thresholds.

In the future, it is planned to explore in more detail and for different datasets as the minimum size required. The training set depends on the total size of the data set and the required accuracy of the model.

5. Acknowledgements

This work was supported by Hankuk University of Foreign Studies Research Fund.

References

[1] Abdi H 2007 The method of least squares. Encyclopedia of Measurement and Statistics (CA, USA: Thousand Oaks)
[2] Adomavicius G and Tuzhilin A 2005 Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions IEEE Transactions on Knowledge and Data Engineering 17(6) 734–49
[3] Adomavicius G and Tuzhilin A 2011 Context-aware recommender systems. Recommender systems handbook pp 217–53
[4] Demsar J 2006 Statistical comparisons of classifiers over multiple data sets J. Mach. Learn. Res. 7 130
[5] Dietterich T 2000 Ensemble methods in machine learning Lect. Notes Comput. Sci. 1857 115
[6] Dunn O 1961 Multiple comparisons among means Am Stat. Assoc. 56 52–64
[7] Dzeroski S and Zenko B 2004 Is combining classifiers with stacking better than selecting the best one? Mach. Learn. 54 255–73
[8] Freund Y and Schapire R 1996 Experiments with a new boosting algorithm Proceedings of the 13th international conference on machine learning pp 148–56
[9] Friedman M 1940 A comparison of alternative tests of significance for the problem of m rankings Ann. Math. Stat. 11 56–92
[10] Goldberg D, Nichols D, Oki B M and Terry D 1992 Using collaborative filtering to weave an information tapestry Communications of the ACM 35(12) 61–70
[11] Gorodetsky V, Samoylov V and Serebryakov S 2010 Ontology-based context dependent personalization technology Web Intelligence and Intelligent Agent Technology (WI-IAT) 2010 IEEE/WIC/ACM International Conference vol 3 Aug pp 278–83
[12] Hariri N, Castro-Herrera C, Mirakhorli M, Cleland-Huang J and Mobasher B 2013 Supporting domain analysis through mining and recommending features from online product listings Software Engineering IEEE Transactions 39(12) 1736–52
[13] Hand D and Vinciotti V 2003 Local versus global models for classification problems: fitting models where it matters Am Stat. 57(2) 124–31
[14] Lopatin E 2019 Methodological Approaches to Research Resource Saving Industrial Enterprises Int. J. of Energy Economics and Policy 9(4) 181-187
[15] Meynkhard A 2019 Energy Efficient Development Model for Regions of the Russian Federation: Evidence of Crypto Mining Int. J. of Energy Economics and Policy 9(4) 16-21
[16] Mikhailov A 2018 Pricing in Oil Market and Using Probit Model for Analysis of Stock Market Effects Int. J. of Energy Economics and Policy 4 43–53
[17] Mikhailov A 2018 Volatility Spillover Effect between Stock and Exchange Rate in Oil Exporting Countries Int. J. of Energy Economics and Policy 8(3) 321–6
[18] Morozko N, Morozko N and Didenko V 2018 Unbalanced liquidity management evaluation of the Russian banking sector J. of Reviews on Global Economics 7 487-496 DOI: https://doi.org/10.6000/1929-7092.2018.07.45
[19] Morozko N, Morozko N and Didenko V 2018 Modeling the process of financing small organizations J. of Reviews on Global Economics 7 774-783 DOI: https://doi.org/10.6000/1929-7092.2018.07.74