International Journal of Intelligent Information and Management Science
ISSN: 2307-0692, Volume 6, Issue 4, August, 2017

Research on the Challenges of Machine Learning

Xiaohan SUN[1*], Kai WANG[2]
[1]WeiFang NO.1 Middle School, Shandong, 261205, China
[2]Baidu, Beijing, 100089, China

Abstract: This study starts from the perspective of the challenges of machine learning, mainly explores the high-dimensional spatial structure and the overall data, and analyzes with the number of calculated data to find the optimal solution. And then, this paper explores combined with the current hot topic that people concerned about to make people put more attention on the machine learning.

Keywords: Machine learning; Deep learning; Probabilistic graphical model; Big data; Optimization problem

1. Introduction

With the development of the times and the continuous progress of science and technology, machine learning has been rapidly developed and utilized; it plays an increasingly widely used role in the social production. But at the same time, some of the problems in machine learning have gradually emerged and became the focus of the times. We need to combine the systematic architecture of machine learning to conduct in-depth exploration to provide more helps for machine learning in the application of social life.

2. Challenges of Machine Learning

From the perspective of rapid development of the society, machine learning has made great progress in China, but at the same time a lot of problems are gradually revealed. The following study mainly focuses on some of the problems with the actual situations.

2.1. Lack of high dimensionality space and samples

In the case of some practical problems, in general, the characteristic function values are high. For example, in the case of graphics processing, taking samples of some single characteristic value, it will find that its characteristic dimension is large; while taking samples of the other characteristic value, the high dimensional spatial features will be found. In organizing the text structures, the functional meaning should be confirmed and through the performance of characteristic dimension to allocate high dimensional space.

The following content takes the samples' characteristic values of high dimensional space into account, in the case of conceptual density distribution, some of the ranges with samples' characteristic must be measured to obtain a specific parameter value, or the type of data can be divided by the relevant estimation result. Different network structures have a high dimensional density probability distribution, so the number of samples in the two-dimensional space should be measured in conjunction with the function value. The number of samples in the same space increases as the dimension increases, and this phenomenon can not be ignored in the analysis of the samples. Since the number of samples can not reach the predetermined requirement when estimating the dimension, in graphic recognition process it can not be divided by simple text classification, so the structure of high dimension will lead to difficult data analysis. The corresponding knowledge training can be obtained through appropriate classification arrangements so as to reduce the pressure of graphics operations by dimension reduction. Or more training data with distinctive characteristics can be obtained by comparing separate sample data and combining random variable structure under the independent conditions. While providing important help for random variables and sample storage. If there is a separate data model among random variables, there is a pressure on data operations when establishing separate analysis of the samples. Therefore, combined with the actual situation for analysis, it can be found that the separate high dimensional data model can only put forward operation of conceptual structure by embedding in low dimensional samples. In the case of such analysis and comparison, the probability model should be calculated with concept of data. At the same time, the characteristics of the sample structure show that in the conduct of comprehensive learning and thinking, problems should be solved by the dimension measurement.

2.2. Difficulty of finding the optimal solution

For the time being, most of the machines still use the method of transforming objective function to obtain complex optimal solutions. This is very useful for the target sample function. In the determination of the samples' function characteristic values, it is always influenced by its constant and linear spatial structure, but it

can be optimized by formalized variable structure. But in general, it is not possible to obtain the range of optimal solution by understanding the change of the objective function value. In addition, in obtaining the optimal solution by optimization techniques, some optimization techniques may consume too much cost, and the final range of parameter values can not meet the corresponding requirements. In machine learning, the problem of some machines can be understood by the release of the optimal solution. By mastering and understanding the learning situation of the machine, the optimal algorithm that is the most suitable for the optimal solution should be designed. This is very important for the current machine learning area. In the calculation of local measurement of the unit area, it should be measured by the initial value of the algorithm to obtain different optimal solution. Or the global variables can be found through the range of the machine learning field. In general, people obtain the optimal solution only through a simple local decomposition method for the calculation of the initial value in local. If people get different algorithm initial values, in the calculation of the function there will always be different variables.

Therefore, the calculation of the global optimal solution is particularly difficult, more parameter variables can only be obtained through the gradient calculation method, and more local variables can be obtained by the optimized function space and the measured scope. After the corresponding algorithm cognition, the corresponding optimal solution range can be obtained. If these optimal solutions do not get corresponding parameter variable in the measurement of local extremum, the objective function is only a kind of corresponding function in the construction of the corresponding problem. Through the optimized functional problem transformation can be measured again. However, it is not clear that if the optimal solution for this objective function must be obtained by this way, or whether the optimal solution can be obtained by this way.

2.3. Poor interpretation

Through the above research, it can be seen that in the function modeling of machine learning field, the actual situation should be taken into consideration for the research and analysis to build the corresponding module, which will achieve the function of machine learning. Then in machine learning, the poor interpretation may be also linked to the construction of practical problems. In the output of the corresponding text, different areas should be controlled and combined with the model structure of the sample to spread. Or in the output of some molecular material, the analysis of corresponding function model can be obtained through the analysis of the sample structure, and the corresponding data can be monitored. In the organization of data prediction, the machine learning also need a comprehensive structural data construction, or through the conversion of corresponding model parameters, the poor interpretation can be explained. On the basis of satisfying the range of output and input, extensive explanations can be obtained by constructing good machine learning differences. In the calculation of optimal solution model, more data can be obtained by analyzing the corresponding questionable data.

From the perspective of current numerical variables and the corresponding relationship determination, if there is no measurement within causal ranges, in a scientifically and rationally systemic construction, the concept of causal relationship need to be combined to get analysis results with time characteristics. Besides, on the basis of the construction of function model, with the corresponding content of the time of big data to analyze the relevant causal relationship, or with the scientific and effective conceptual identification to carry out the analysis, so that more explanatory results can be obtained

2.4. Calculation of large data quantity

In the measurement of data quantity, it must be combined with the characteristics of the time for analysis. The following study is mainly focus on the application of data quantity, or through the measurement that in the field of large data quantity and the construction of the basic parameters in scope of machine learning, it is very important for the calculation of the current large data quantity. The requirements of the application calculation of large data quantity, algorithm structure that is more adapt to the current machine learning, and the use of high performance corresponding parameter analysis should be taken into account. Besides, debugging in accordance with the working state and combined with the methods of basic theoretical study can be used to solve practical problems. However, the development of machine learning is very fast for the overall parameter structure. Many people have integrated different theoretical basic knowledge learning and overall as a whole, or through the exploration of practical problems to obtain the calculation of large data quantity.

At present, due to the large amount of large data quantity calculation, in the development of corresponding algorithm and analysis structure, it must be allocated combining with the learning of big data information. Because the development of big data is very rapid, it is difficult to form a coherent in complexity of time and space. The post-training large data quantity code structure has to obtain the corresponding structure by repeated operation, or use the method of learning adjustment to complete the calculation of parameters. In support of the huge data uniform, analyzing combined with the machine learning algorithm to get more relevant variable structures. On the basis of rapid equivalent operation, it should accurately obtain the size of certain values by searching and analyz-

ing the proportion of the corresponding parameters. However, in the machine learning process, it is difficult to analyze by operation structure of the algorithm, but can be dealt with by combining with the accurate algorithm precision operation.

3. Several Important Questions of Machine Learning

3.1. Big data

In the process of machine learning, it is easy to meet some big data problems. It is mainly because of the data storage and conversion through the application of Internet technology, which is very important for the current information processing. The learning of big data in machine learning is always difficult to sum up and study because of its large amount. These synthetic algorithms, which are processed by data distribution, often have some unique architecture. At this time, there are always a lot of difficulties in the collating and analyzing learning algorithms. If build the corresponding parameter system simply in learning process, it is difficult to complete the design of machine learning algorithm. In the analysis and study of data, it is necessary to obtain the data with basic structure through the distribution of different data quantities. Or for the application of large amounts of data, it can find a small amount of data through some phenomena, which is not yet found in the past. In the corresponding application relationship, optimize the algorithm structure of big data, so that the structural model can obtain a more complete performance system, through the analysis of the algorithm data to get more practical results. Reducing the drawbacks of fewer samples through the standardized operation is of great significance for the calculation of sample size. For the study of structural models under different systems, the classified physical structure basis is easy to be compared and calculated. It has a certain help for the current regional structure calculation by combining with changes proportion of parameters of the big data and obtaining more performance parameters in the structuring model. In the era of big data information, if the overall algorithm structure changes, it will lead to the measurement of samples' number affected. The basic identification of big data information provides some compensation for the algorithm with different structures.

3.2. Deep learning

In the process of deep learning, it is mainly rely on artificial neural network for depth measurement. More effective function knowledge system can be obtained through the reverse transmission of data information. On the basis of constructing the perfect neural network structure, the learning algorithm and the basic unit are built through the different representation of continuous function. For the neural network, if there are some local extreme value, it will lead to that the objective function is difficult to normally achieve. The limitation of the sample size is also obtained by exploration of learning. In the process of deep learning, it is necessary to gain a deep understanding of the implicit phenomena of large numbers of data samples. Through the comparison of different parameters under the neural network to carry out effective data monitoring, this can be simply paid to get more data information. For example, in the processing of graphics information, more effective information can be obtained through deep calculation and combining with specific measurement methods for processing. Then the deep learning must optimize its generalization ability.

4. Conclusion

With the development of the times, the progress of science and technology, and the advent of the information age, machine learning has brought more and more benefits to social production. This paper explores with challenges of machine learning, hoping to provide more help for the future machine learning.

References

[1] Zhang Run, Wang Yongbin. Research on Machine Learning with Algorithm and Development [J]. Journal of Communication University of China Science and Technology, 2016,23(02):10-18+24.

[2] Huang Yihua. Research Progress of Large Data Machine Learning System [J]. Big Data, 2015, 1(01):35-54.

[3] Zhang Changshui. Changes of Machine Learning [J]. Chinese Science: Information Science, 2013,43(12):1612-1623.