ROOTINDEXING
JOURNAL ABSTRACTING AND INDEXING SERVICE
INTERNATIONAL ADVANCE JOURNAL OF ENGINEERING, SCIENCE AND MANAGEMENT (IAJESM)
July-December 2023, Submitted in July 2023, iajesm2014@gmail.com, ISSN -2393-8048
Multidisciplinary Indexed/Peer Reviewed Journal. SJIF Impact Factor 2023 =6.753

A Study on the Research and Development of Machine Learning Algorithms

Sabale Swarupa Nivrutti, Research Scholar, Shri JJT University Jhunjhunu, Rajasthan
Dr. Hiren Dand, Research Guide, Department of Computer Science and Engineering, Shri JJT University Jhunjhunu, Rajasthan

Abstract

The fundamental categories of machine learning—supervised, unsupervised, and reinforcement learning—are examined in this article. It incorporates study of popular machine learning methods, including boosting and bagging, BP, artificial neural network, random forest, decision tree, and SVM algorithms. The goal is to raise public awareness of machine learning and quicken its rate of popularization through the creation of theoretical systems, enhanced autonomous learning capabilities, integration of various digital technologies, and promotion of tailored custom services.

Key words: Machine Learning, Development, computer technology

Introduction

The quick advancement of science and technology has given rise to new development opportunities, one of which is artificial intelligence. The incorporation of multidisciplinary theoretical knowledge, such as statistics and algorithm complexity, into machine technology, which is based on computer technology, enhances the functional characteristics of artificial intelligence. It is possible to improve the applicability of machine learning algorithms and add convenience for the industry's economic development by conducting a realistic analysis of these algorithms and offering direction reference for later machine learning development.

Machine Learning Classification

Supervised Learning

Supervised learning is one of the more fundamental learning techniques used in machine learning. This approach to learning describes how individuals set appropriate learning objectives prior to learning. When the machine is first trained, it uses information technology to understand what it needs to learn. We are expected to progressively finish the necessary learning content in a supervised setting in order to gather fundamental data. In contrast to alternative learning approaches, supervised learning is capable of fully utilizing the machine's capacity for generalized learning. Once the system has finished learning, it can assist users in resolving some highly systematic problems related to regression or classification. Currently, BN, SVN, KNN, and other traditional learning techniques are frequently employed. The machine learning process exhibits a certain regularity and the learning content is more systematic because the entire learning process has a purpose.

Unsupervised Learning

Unsupervised learning is correlated with supervised learning. In what is known as "unsupervised learning," the machine analyzes the data on its own, without the need for human intervention, by marking the content in a certain way during the learning process. In actuality, the technique of operation involves first allowing the machine to learn the fundamental ideas and material, and then granting it sufficient autonomy to finish a sequence of content learning tasks that involve ideas and content resembling the fundamental principles, like tree roots. The scope of machine learning content has generally risen due to the gradual improvement of learning. Deep belief networks and autoencoders are two examples of techniques used in unsupervised learning nowadays. These conditions are favorable for solving clustering issues and have useful applications in the growth of numerous enterprises.

Reinforcement Learning

There are machine learning applications of reinforcement learning in addition to supervised and unsupervised learning. The methodical learning of a given subject is known as "reinforcement learning." The information gathered over the prior time frame will be utilized in the particular application procedure. It creates a closed loop of data processing by

organizing and processing the feedback data of a specific part. All things considered, dynamic learning and statistical data collection are expanded upon by reinforcement learning. These techniques are mostly applied to robot control issues. The Q-learning algorithm and the temporal difference learning algorithm are two of its representative learning techniques.

Frequently Analysis used Machine Learning Algorithms

Decision Tree Algorithm

The decision tree algorithm is one of the popular machine learning algorithms that is a part of the traditional algorithm content. Its basic idea is that data information is processed beginning at the collection instance's root node and ending at the point where the nodes converge to form the complete data set. Practical examples divided scientifically. The decision number method will keep splitting branches to make data analysis easier, and at the same time, it will trim the branches to enhance the data content's integrity. The algorithm falls under the top-down algorithm from a computational perspective. The node's content is examined for the best attributes throughout the content analysis process, and the node is then enlarged to include more than two based attributes. This allows you to obtain detailed information about the split. Additionally, by using a branching method akin to a tree, you can increase the number of samples that are available for analysis and, based on sample number statistics, identify the content that has the greatest number of samples in the classification. For instance, you can define the highest limit of branch splitting and designate the decision tree with a lot of data as tree A while evaluating data. If the upper limit is set at 5, then the larger tree A will be classified as having reached that value. It will then cease splitting and process the larger tree model using the pruning strategy, which will help to improve the data quality and scientific validity of the analysis.

Random Forest Algorithm

Like the decision tree technique, the random forest algorithm can be used to process data further during data calculation. In the course of practical use, the random forest method will be helpful in managing irrational data. hence enhancing the data split findings' scientific validity and the data analysis results' accuracy. Multiple sets of classification trees will be built simultaneously during the data analysis process, and regression processing will be handled by the unified method. Assuming the decision tree is an independent set ai (i= 1,2,3 ... n), then the random forest is the total set A, where A = {a1, a2, a3, ..., an}, where a = 1,2,3 ...

n. Each set remains independent, and the distribution is a state of random distribution. When evaluating the classification data information, it will be selected by means of voting. The classification with the highest number of votes in the voting will output the vector value xi, and then the vector content will be classified to calculate the average value of different score states and provide data reference for the final judgment.

Artificial Neural Network Algorithm

In order to do complicated memory tasks, an artificial neural network (ANNN) mimics the way that information is sent by humans by classifying various inputs into a single neuron and linking those neurons via the Internet. Nonetheless, this developing data analysis procedure serves as the foundation for the artificial neural network method. Every digital unit among the distinguished neurons has a high level of authenticity, and the data can finish the external output process. It is analogous to how the human body advances, pauses, and then sprints. The data information supplied in the artificial neural network method has a wide range of application features, and the associated analysis process can be finished in accordance with actual needs. Currently, self-organizing neural networks (SOM), ART, and multilayer forward neural networks (MLFN) are among the frequently utilized artificial neural networks. We can pre-set the weighting coefficient and output threshold to make the data analysis and computation process easier. The orderliness of the entire numerical analysis process is improved when a certain value is output to the outside once the calculated sum surpasses this amount.

SVM Algorithm

The SVM algorithm is among the frequently utilized algorithm content in machine learning. The algorithm mostly uses the vector machine approach in the particular application procedure to finish the predetermined data analysis work. In order to optimize the data information, the SVM algorithm will simultaneously examine the data to be processed using the automatic support of the SVM. To enhance the scientific quality of the final data analysis results, it is necessary to gather numerous sets of analysis samples during the analysis process in order to identify the sample data of the boundary value. For instance, if the data to be processed is H (d), the first step in processing it would be to process it centrally using SVM technology, allowing the data to be fully distributed. Second, the maximum distance of the entire plane is used to calculate the boundary of the H (d) plane. Ultimately, the output vector is obtained by analyzing the vector content of the H (d) plane, hence enhancing the precision of data processing.

Boosting and Bagging Algorithms

The main application benefit of the boosting algorithm, a new kind of machine algorithm content, is that it can finish the accurate processing of data information and raise the accuracy of the processing result in the end. In actuality, the Boosting algorithm will be used to build the function prediction system, and the reinforcement learning mode will be used to continuously optimize the system content, speeding up the processing of data. One of the Boosting algorithm's more basic applications is AdaBoost. AdaBoost is a significant assurance for the Boosting algorithm's growth at the same time. The data processing technique is highly similar to the Bagging algorithm. The training set is chosen at random by the Bagging algorithm in the real application. Additionally, the Bagging method does not assess the weight content when calculating the function model; therefore, in order to increase the accuracy of the data analysis findings, we must continuously modify the data model with the use of training.

BP Algorithm

The supervised learning category includes the BP algorithm. Figure 1 illustrates the algorithm's fundamental idea. A shallow forward neural network computing model with an input layer, a hidden layer, and an output layer is depicted in the image. As network nodes, a vast number of neurons are linked to one another. Via an excitation function, each neuron interprets the connection strength signals as network weights. The pattern information in the input data is transferred to the output layer by varying these connection strengths.

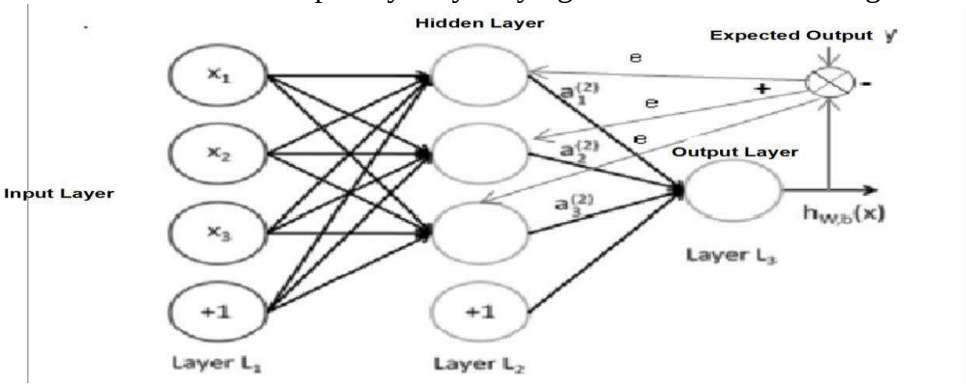

Figure 1 Basic Principles of Algorithm Application

As can be seen from the above figure, the direction of the information flow of forward propagationis input layer → hidden layer → output layer, and its mathematical model is:

$$h_{W,b}(x) = f\left(\sum_{i=1}^{n} W_i x_i + b\right)$$

Where Wi and b are their weights and bias parameters, f (W, b; x): R → R is called the excitation function, and sig-moid can be selected in practical applications, Tanh, ReLU and other functions or their variants, $h_{W,b}(x)$ are the network output values.

In practical applications, the BP algorithm can be implemented by the steepest descent method, Newton method and its improved algorithm, quasi-Newton method and its correction algorithm, etc. At present, the L-BFGS algorithm is most widely used, and non-precise line search methods are often used to complete the optimization. This method follows Wolfe's criterion and Armijo's criterion, which guarantees the balance between the decline of the cost function and the convergence of the iterative sequence.

Research on Machine Learning Development

Theoretical System Continues to Mature

The mechanical theory system will undergo further optimization in the coming development phase, along with an expansion of its content branches and coverage. When machine learning content is first being developed, it is mostly suitable to specific automation businesses, and the theoretical system's content is not entirely sound. Certain fields do not allow the substance of their theoretical framework to be used in practice. In reaction to these circumstances, the theory of machine learning will advance steadily, and the level of content refinement will also advance, creating favorable conditions for the eventual advancement of machine learning.

Autonomous Learning Ability is Further Improved

Many Chinese businesses have adopted the automation development model as of late, and the next phase of development will center on intelligence. The ability of machines to learn on their own will be significantly enhanced in light of the Internet's rapid expansion. The autonomy that machine learning may achieve will keep growing, regardless of whether it is supervised or unsupervised learning. In the machine's future learning process, it will learn in a targeted or comprehensive way based on its own requirements. This lowers the cost to the business of updating the equipment structure and lays the groundwork for the enterprise economy to grow steadily.

Integration of Multiple Digital Technologies

Currently, relying solely on Internet technology has resulted in the development of numerous ancillary technologies, including digital technology, cloud computing, Internet of Things, and so forth. During the data calculating process, these technologies can offer a variety of practical conditions. Even if the integration of various digital technologies is still in its early stages, it is constantly improving due to the rapid advancement of technology. Furthermore, these technologies will be mixed with algorithms in a future development process to create a new technological application system, setting the groundwork for future advancements in data analysis speed.

Promotion of Personalized Customization Services

People's needs for tailored applications are consistently growing along with the socioeconomic level, and this is one of the key future development directions for machine learning. As the intelligent mechanical learning level continues to advance, various application modules can be configured to meet the real-world requirements of users. In order to satisfy the user's individual demands and raise user satisfaction with services, the data module can simultaneously match the related service content and filter out the corresponding information content after receiving the user request message.

Conclusion

In conclusion, machine learning is still in its infancy, mostly dependent on supervised learning, and unable to completely replace artificial intelligence that is not very strong. The theoretical underpinnings and practical applications of machine learning require ongoing improvement by pertinent experts. We should provide a favorable atmosphere for machine learning in the related scientific field and in the advancement of computer technology, as machine learning has a very bright future. Setting up machine algorithms appropriate for the growth of local businesses, actively learning from the experiences and lessons of developed nations, and offering technical assistance for the industry's economic development are also essential.\

REFERENCES

[1] Jiang Na, Yang Haiyan, Gu Qingchuan, Huang Jiya. Machine learning and its algorithm and development analysis [J]. Information and Computer Science (Theoretical Edition), 2019 (01): 83-84 + 87.

[2] Li Kanghua, Jiang Shan. Machine Learning and Cultural Production Reform——Based on the Perspective of the Development of AI Technology [J]. Journal of Xiangtan University(Philosophy and Social Sciences), 2020, 44 (01): 74-79.

[3] Li Zhiwei. Development of machine learning and several learning methods [J]. Industry and Science Forum, 2018, 15 (10): 198-199.

[4] Zhang Changshui. Research on the development of machine learning and data mining [C].. 2010-2011 Development Report of Control Science and Engineering Discipline.: Chinese Society of Automation, 2018: 82-89 + 223.

[5] Zhang Run, Wang Yongbin. Research on machine learning and its algorithm and development [J]. Journal of Communication University of China (Natural Science Edition), 2018, 23 (02): 10-18 + 24.