Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/36/20230425

Current study on human-computer interaction in machine learning

Yalan Cai[1], Yuyang Lu[2, 3]

[1]Fuzhou No. 8 High School, Fuzhou, China
[2]Nanjing No.13 High School International Department, Nanjing, China

[3]1910751103@mail.sit.edu.cn

Abstract. Machine learning has become one of the research hotspots at home and internationally due to the continued growth of artificial intelligence, and the application of machine learning is more and more widely developed. In the process of applying machine learning methods to real problems, there are defects that lead to biased results. This paper discusses the importance and necessity of human-machine interaction in the application of machine learning methods, as well as where human-machine interaction occurs, and puts forward two questions: "whether human should interact with machine in the process of machine learning" and "how to make machine learning have better performance". To answer the above two questions, this paper concludes that in the application of machine learning methods, people with certain professional knowledge can get better results in the machine learning process. Further, when machine learning is applied to the real world, there are some flaws that lead to failure or unsatisfactory results, and this paper proposes a way to improve this undesirable phenomenon by involving people in the machine learning process. Finally, this paper summarizes the main shortcomings of current machine learning, clarifies the development direction of machine learning that must be anthropocentric, and expresses some views on machine learning.

Keywords: interactive machine learning, human-computer interaction, anthropocentric development.

1. Introduction

Machine learning has made remarkable achievements in many fields, resulting in many machine learning applications. Machine learning is a tool for extracting insightful knowledge from sample data, and its success lies in its ability to learn from massive amounts of data [1]. Different machine learning methods are roughly partitioned into supervised learning and semi-supervised learning, but both have advantages and disadvantages. This paper points out a machine learning method that incorporates supervised learning and semi-supervised learning, which can greatly solve the disadvantages of these two machine learning methods. The current machine learning methods have exposed some defects in the application, which cause that the results cannot produce satisfactory data for users every time [2]. This paper guesses that this is due to the lack of sufficiently high-quality training models, which, however, cannot be accomplished without reliable information sets. Finding a solution to this issue is the main goal of this essay. Much of the literature on machine learning does not clearly define human-machine interaction in machine learning, or ignores the importance of humans. This paper compared the

DOI: 10.54254/2755-2721/36/20230425

two results obtained with or without human-machine interaction, and concluded that in previous machine learning applications, the absence of human participation is an important factor leading to the bias of results. After investigation, it is found that the output results in machine learning applications are not satisfactory, and human experience and knowledge can be used to make up for the shortcomings, so machine learning ignores the shortcomings of human circulation. Therefore, this paper believes that in the process of machine learning applications in the future, it must be anthropocentric to get more accurate results. According to the current situation, this paper analyzes a series of future development directions of machine learning, and finds that machine learning lacks the function of extracting important information, and a large amount of useless information will have a negative impact on work efficiency. We believe that improving the user's human-computer interaction experience is a more convenient way to obtain more accurate data. The applications are similar across all domains and can help machine learning get better results across the board.

2. Analysis of machine learning methods

2.1. Supervised learning

In supervised learning, the tutor or teacher provides the objective Y, which is to assign a label to each example indicating what the machine learning algorithm should learn to do. Representative algorithms include deep neural networks, decision trees, logistic regression, and linear regression, etc. Supervised learning application example: obtaining labels directly or indirectly from data through manual labelling or user feedback. An example is text classification. A batch of texts is first manually annotated and then a model is trained in a supervised manner to classify the unknown texts by topic [3].

2.2. Unsupervised learning

As opposed to unsupervised learning, there is no tutor or teacher, that is, the samples are not labelled, and the algorithm must learn hidden patterns in the data without guidance. Representative algorithms are clustering algorithms, such as K-Means, and density estimation algorithms. Unsupervised learning application example: User clustering. Birds of a feather flock together, and similar people will behave similarly. For example, we can cluster users based on their behavior on the website to achieve more accurate user profiles [4]. Figure 1 depicts the relationship between the various machine learning methods.

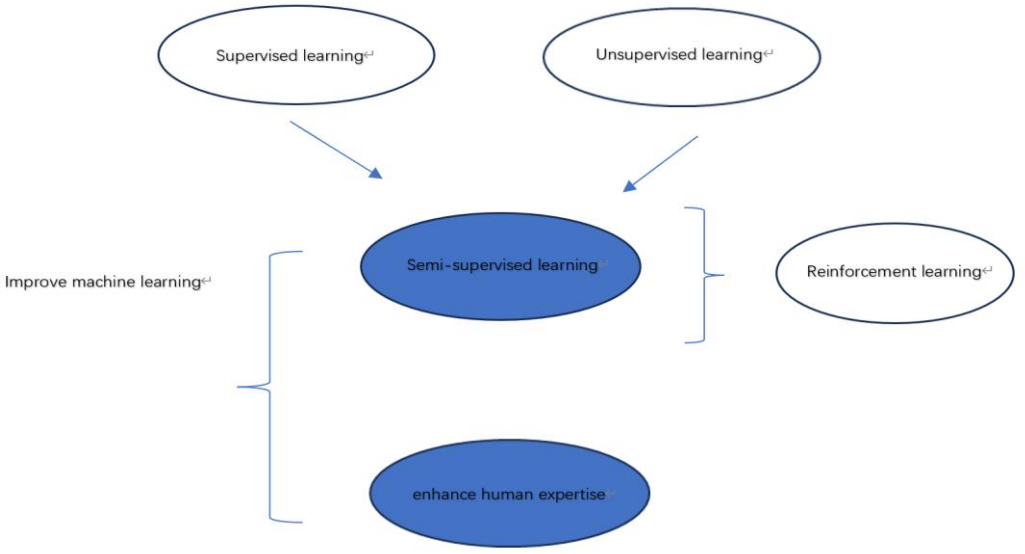

Figure 1. Several machine learning methods.

DOI: 10.54254/2755-2721/36/20230425

3. The difference between machine learning with human and without human

3.1. Approach to human-computer interaction without machine learning

In Florian Kunneman's experiments on vaccination, they needed a way to combine the use of much larger amounts of training data with the predictive feedback provided by human-computer interaction systems [5]. In the process of solving complex problems in reality, there will still be unsatisfactory results in machine learning. The way to avoid this is to let the human-machine interaction in the machine learning process to get more accurate results.

3.2. Approach to human-computer interaction with machine learning

In Changyu Deng's paper, they found that high-precision data was difficult to obtain or expensive. However, integrating human knowledge into it can enable capabilities that machine learning currently cannot and facilitate human-machine interaction between humans and machines [6]. In machine learning, using the method of human-computer interaction, the original difficult and expensive difficulty of obtaining data is solved. Machine learning lacks the superior abstract reasoning skills that humans possess, together with a wealth of prior experience and information. Especially when introducing new patterns and complex processes. While machine learning techniques can only deliver clear and accurate answers to questions that are well-structured, they cannot produce high-quality responses to ambiguous questions. From the above comparison, it can be seen that human participation is vital in machine learning.

4. Ways to improve machine learning

4.1. Better machine learning methods

4.1.1. Semi-supervised learning. The two types of learning techniques based on data mining are supervised learning and unsupervised learning. Supervised learning is an optimization process, which is dedicated to constructing the optimal model to minimize the error between it and the known knowledge under the premise of the labelled training sample set. The goal is single and clear, and the accuracy is high. Unsupervised learning does not have an exact target, but it can seek meaningful patterns according to the characteristics of the object without labelled training samples, but the accuracy is relatively low [7]. This paper found that the combination of supervised and unsupervised learning was the most commonly used learning method and the one with the most stringent data obtained.

The process of semi-supervised learning is very rigorous. In the first step, train the model with a small amount of labelled data. In the second step, the trained model is used to predict the unlabelled data. The third step selects samples with high label certainty, adds training samples, and retrains the model. Finally, the second and third steps are repeated until the model effect converges. Data preprocessing is done before applying machine learning, if the result is too different from the preprocessed result, the quality is considered insufficient and the loop continues until a result comparable to the preprocessed result is obtained [8].

4.1.2. Reinforcement learning. Reinforcement learning is one kind of semi-supervised learning. To maximize a monetary incentive signal, it is learning how to translate states into actions. The agent tries different actions without guidance to discover which action brings the most reward. Representative algorithms include Q-Learning, deep Q-network, etc. Reinforcement Learning Example Application: Robot Control. Robots learn by trial and error through rapid interaction with the environment, such as Boston Dynamics learning to dance. Through constant dance movements, the robot starts again after falling. After repeated attempts, it can gradually master the skills of dance [9]. Figure 2 illustrates the process of semi-supervised learning to get the training model.

DOI: 10.54254/2755-2721/36/20230425

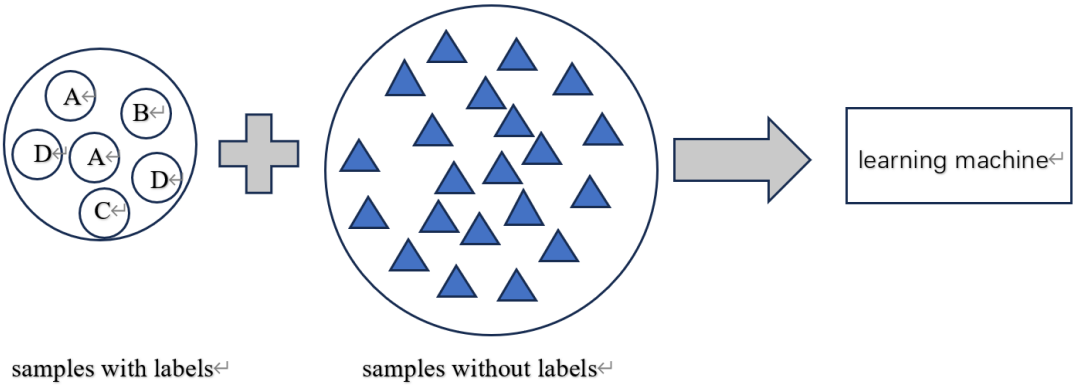

Figure 2. Semi-supervised learning is the least flawed machine learning method [7, 8].

4.2. Improve the professional of human who interact with machine

In the application of machine learning, improving the expertise of the human interacting with the machine is another way to improve the performance of machine learning [10].

In the process of applying machine learning, the human interacting with it does not need to have a deep knowledge of machine learning methods. Instead, the level of domain knowledge needed by a human expert varies on the machine learning approach, the field of application, the job, and task complexity of the person in the loop. When interacting with machine learning techniques, people's level of topic expertise might be crucial. Additionally, more sophisticated expertise is needed for activities with a higher level of complexity in human-machine method interaction.

In general, for people who collaborate using machine learning methods, higher levels of expertise lead to better human-computer interaction outputs. Domain specialists, however, are sometimes substituted by crowd workers or individuals who lack knowledge in specific machine learning applications due to other factors like cost and time [11]. The research examines human performance at various machine learning application levels, and the results support the need for people to have domain expertise. The outcomes demonstrate that domain specialists outperform non-specialists in data annotation.

5. Discussion

5.1. Problems of current machine learning based on human-computer interaction

5.1.1. Information overload. Machine learning lacks the function of filtering important information. Collecting and saving a large amount of data is one of the main working methods of current artificial intelligence in the field of human-computer interaction, but the accumulation of a large amount of irrelevant data will cause some negative effects, such as excessive energy consumption and information overload, and a large amount of irrelevant data will make people go to spend a lot of time to filter the data that need to be used, reducing the efficiency of people's work.

5.1.2. Limitations in solving problem. In the current human-computer interaction, machine learning methods have certain limitations. Initially, in order to achieve automation and objectivity, machine learning methods do not model people. As a result, the ability to solve broader problems is sacrificed. Machine learning methods are limited to solving clearly specified problems with existing human knowledge, and this lagging approach makes it difficult to solve significant problems that exceed human knowledge [12].

DOI: 10.54254/2755-2721/36/20230425

5.1.3. Uncertainty in classification. When machine learning is applied in the field of human-computer interaction, it is common to collect a large amount of data and then categorize it. Most of the traditional machine learning methods using microarray data have uncertainty and lack sufficient robustness in classification. The collection process of microarray data often adds noise sources and usually has limited selection of data. It makes the classification full of instability [13].

5.2. Current approaches that can solve the problems of machine learning in human-computer interaction

5.2.1. Machine learning systems that can predict user relevance. Facing the problem of inefficiency in processing information caused by a large amount of irrelevant information, this problem can be avoided by developing machine learning systems that can predict the relevance of the user in order to selectively present data to the user. The work of machine learning systems is divided into two main stages (1) Capture the information seeking behavior of past users in performing tasks to statistically determine the type of relevant data and then model it. (2) Highlighting the relevant data to the user through these models. Through data training, this machine learning system can help users to find information more efficiently [14].

5.2.2. Anthropocentric machine learning systems. In order to break through the limitations of machine learning methods in human-computer interaction, the machine learning and modelling process should be compatible with humans. In the process of solving this problem, machine learning system needs to model the user and the computer will take the role of an assisting helper to give suggestions to the user iteratively and then the user will make new future decisions based on the feedback [12]. The human will be directly at the centre of the machine learning system by being part of the machine learning. This anthropocentric interactive machine learning approach will help people solve a wider range of problems.

5.2.3. Classification algorithms based on integrated methods. In order to eliminate the uncertainty associated with microarray data, a machine learning approach based on integrated learning methods can be developed. Integrated learning methods use an approach that combines a set of different base classifiers to make more accurate classification decisions. The method is divided into three main steps (1) sampling of base data (2) generation of alternative base classifiers (3) ensemble of classifiers. The classification performance of integrated learning method is more general, accurate and has higher robustness than other classification methods [13].

5.3. Future Future directions for machine learning related to human-computer interaction

The future development of machine learning will aim at improving user experience. In the field of medical orthopedics, for example, advances in machine learning methods have made it possible for even non-computer literate doctors to use AI fluently. In the field of orthopedics, machine learning is utilized to assist physicians in making predictions that can be used to direct and anticipate the outcome of surgery. In order to assist doctors in their following diagnostic work, related algorithms can be utilized to analyze the risk of arthritis in patients with fractures [15]. Not only has machine learning undergone a difficult evolution in orthopedics, but it has also undergone a similar development in other domains. Machine learning techniques in the field of human-computer interaction will progressively become more anthropocentric, assisting more industries in realizing the simplicity of the process, automating it, and creating a connection between the user and the computer to enhance the user's experience at work in all respects.

6. Conclusion

This paper's research reveals that anthropocentric machine learning approaches are more beneficial and long-lasting than those that have completely removed the human element. In order to make the human an integral element of the machine learning system and the machine into a tool to aid in the development

DOI: 10.54254/2755-2721/36/20230425

of the human, the development of machine learning in human-computer interaction should be anthropocentric. As a result, the anthropocentric objective will continue to guide the development of machine learning techniques in human-computer interaction. The present state of machine learning in the area of human-computer interaction is summarized in this article, which will help other researchers who are interested in the subject matter to continue their research. Current research still has some shortcomings, such as the opacity of machine learning in human-computer interaction that remains unresolved. In further future studies, researchers will actively explore to solve these problems.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Barnova, K., Mikolasova, M., Kahankova, R. V., Jaros, R., Kawala-Sterniuk, A., Snasel, V., Mirjalili, S., Pelc, M., & Martinek, R. (2023). Implementation of artificial intelligence and machine learning-based methods in brain-computer interaction. Computers in biology and medicine, 163, 107135. Advance online publication.

[2] Deng, C., Ji, X., Rainey, C., Zhang, J., & Lu, W. (2020). Integrating Machine Learning with Human Knowledge. iScience, 23(11), 101656.

[3] Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised Machine Learning: A Brief Primer. Behavior therapy, 51(5), 675–687.

[4] Eckhardt, C. M., Madjarova, S. J., Williams, R. J., Ollivier, M., Karlsson, J., Pareek, A., & Nwachukwu, B. U. (2023). Unsupervised machine learning methods and emerging applications in healthcare. Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA, 31(2), 376–381.

[5] Kunneman, F., Lambooij, M., Wong, A., Bosch, A. V. D., & Mollema, L. (2020). Monitoring stance towards vaccination in twitter messages. BMC medical informatics and decision making, 20(1), 33.

[6] Deng, C., Ji, X., Rainey, C., Zhang, J., & Lu, W. (2020). Integrating Machine Learning with Human Knowledge. iScience, 23(11), 101656.

[7] Peng X, Yang J, Hu Y, Li J, & Jia Z. (2008). An early warning method of flavistratus jumping armor based on semi-supervised learning. Agricultural Mechanization Research (3), 4.

[8] Liang J, Gao J, & Chang Y. (2009). Progress in semi-supervised learning. Journal of Shanxi University: Natural Science Edition, 32(4), 7.

[9] Matsuo, Y., LeCun, Y., Sahani, M., Precup, D., Silver, D., Sugiyama, M., Uchibe, E., & Morimoto, J. (2022). Deep learning, reinforcement learning, and world models. Neural networks : the official journal of the International Neural Network Society, 152, 267–275.

[10] Maadi, M., Akbarzadeh Khorshidi, H., & Aickelin, U. (2021). A Review on Human-AI Interaction in Machine Learning and Insights for Medical Applications. International journal of environmental research and public health, 18(4), 2121.

[11] Ni Mhurchu, C., Eyles, H., Jiang, Y., & Blakely, T. (2018). Do nutrition labels influence healthier food choices? Analysis of label viewing behaviour and subsequent food purchases in a labelling intervention trial. Appetite, 121, 360–365.

[12] Çelikok, M. M., Murena, P. A., & Kaski, S. (2023). Modeling needs user modeling. Frontiers in artificial intelligence, 6, 1097891.

[13] Peng Y. (2006). A novel ensemble machine learning for robust microarray data classification. Computers in biology and medicine, 36(6), 553–573.

[14] King, A. J., Cooper, G. F., Clermont, G., Hochheiser, H., Hauskrecht, M., Sittig, D. F., & Visweswaran, S. (2019). Using machine learning to selectively highlight patient information. Journal of biomedical informatics, 100, 103327.

DOI: 10.54254/2755-2721/36/20230425

[15] Martin, R. K., Ley, C., Pareek, A., Groll, A., Tischer, T., & Seil, R. (2022). Artificial intelligence and machine learning: an introduction for orthopaedic surgeons. Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA, 30(2), 361–364.