RESEARCH ARTICLE OPEN ACCESS

Real-Time Machine Learning Algorithms

Yew Kee Wong
JiangXi Normal University, Jiangxi - China

ABSTRACT
There are different ways an algorithm can model a problem based on its interaction with the experience or environment or whatever we want to call the input data. It is popular in machine learning and artificial intelligence textbooks to first consider the learning styles that an algorithm can adopt. There are only a few main learning styles or learning models that an algorithm can have and we'll go through them here with a few examples of algorithms and problem types that they suit. This taxonomy or way of organizing machine learning algorithms is useful because it forces you to think about the roles of the input data and the model preparation process and select one that is the most appropriate for your problem in order to get the best result. Due to the epidemic, online course learning has become a major learning method for students worldwide. Analyzing its massive data from the massive online education platforms becomes a challenge because most learners watch online instructional videos. Thus, analyzing learners' learning behaviors is beneficial to implement personalized online learning strategies with sentiment classification models. The experimental results show that, in the process of "massive data mining," personalized learning strategies of this model can efficiently enhance students' interest in learning and enable different types of students to develop personalized online education learning strategies.
Keywords — Machine Learning, Artificial Intelligence, Autonomy, Cryptography, Privacy Protection

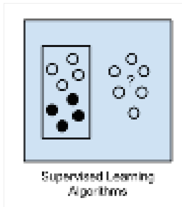
I. INTRODUCTION

When crunching data to model business decisions, you are most typically using supervised and unsupervised learning methods. A hot topic at the moment is semi-supervised learning methods in areas such as image classification where there are large datasets with very few labeled examples. Learner sentiment analysis can be performed by collecting, analyzing, and representing data related to learners' interactions with the course that provides researchers and teachers with an opportunity to understand learners' behavior and assess their performance through their interactions with the video content [5]. Several studies on learner engagement and explorations of patterns of learner behavior with video interactions have focused on analyzing data collected from learners' interactions with different forms of course [6–8].

Educational environments face increasing complexity and diversity [11]; for example, students from different locations can take the same course. With regard to diversity, instructional designers are constantly challenged to adapt to the individual needs of students. Therefore, they must adopt teaching methods that are appropriate for different students [12]. This is the reason for the popularity of personalized learning. Personalized learning defined in the National Educational Technology Initiative supported by the U.S. Department of Education means that learning platforms can optimize learning paths and instructional methods based on the needs of each learner [13]. Such a learning platform allows students to pursue their personal learning goals at their own pace [14]. Thus, the main benefit of personalized learning is its ability to adapt to the needs of different students. This benefit is supported by empirical evidence, which suggests that personalized learning allows instructional designers to meet students' needs and helps students clarify and improve their understanding of learning objectives [15, 16].

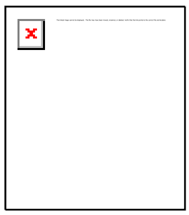
II. DIFFERENT LEARNING STYLES IN MACHINE LEARNING ALGORITHMS

A. *Supervised & Unsupervised Learning*



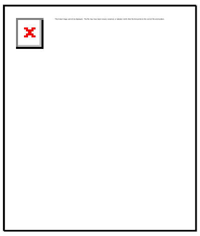
Input data is called training data and has a known label or result such as spam/not-spam or a stock price at a time. A model is prepared through a training process in which it is required to make predictions and is corrected when those predictions are wrong. The training process continues until the model achieves a desired level of accuracy on the training data. Example problems are classification and regression. Example algorithms include: Logistic Regression and the Back Propagation Neural Network. Unsupervised Learning Input data is not labeled and does not have a known result. A model is prepared by deducing structures present in the input data. This may be to extract general rules. It may be through a mathematical process to systematically reduce redundancy, or it may be to organize data by similarity. Example problems are clustering, dimensionality reduction and association rule learning. Example algorithms include: the Apriori algorithm and K-Means.

Input data is called training data and has a known label or result such as spam/not-spam or a stock price at a time. A model is prepared through a training process in which it is required to make predictions



and is corrected when those predictions are wrong. The training process continues until the model achieves a desired level of accuracy on the training data. Example problems are classification and regression. Example algorithms include: Logistic Regression and the Back Propagation Neural Network. Unsupervised Learning Input data is not labeled and does not have a known result. A model is prepared by deducing structures present in the input data. This may be to extract general rules. It may be through a mathematical process to systematically reduce redundancy, or it may be to organize data by similarity. Example problems are clustering, dimensionality reduction and association rule learning. Example algorithms include: the Apriori algorithm and K-Means.

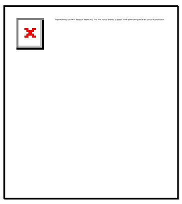
B. Semi-Supervised Learning



Input data is a mixture of labeled and unlabelled examples. There is a desired prediction problem but the model must learn the structures to organize the data as well as make predictions. Example problems are classification and regression. Example algorithms are extensions to other flexible methods that make assumptions about how to model the unlabeled data. Most Common Types of Algorithms

Algorithms are often grouped by similarity in terms of their function (how they work). For example, tree-based methods, and neural network inspired methods. I think this is the most useful way to group algorithms and it is the approach we will use here. This is a useful grouping method, but it is not perfect. There are still algorithms that could just as easily fit into multiple categories like Learning Vector Quantization that is both a neural network inspired method and an instance-based method. There are also categories that have the same name that describe the problem and the class of algorithm such as Regression and Clustering. We could handle these cases by listing algorithms twice or by selecting the group that subjectively is the “best” fit. I like this latter approach of not duplicating algorithms to keep things simple. In this section, we list many of the popular machine learning algorithms grouped the way we think is the most intuitive. The list is not exhaustive in either the groups or the algorithms, but I think it is representative and will be useful to you to get an idea of the lay of the land.

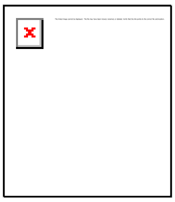
C. Regression Algorithms



Regression is concerned with modeling the relationship between variables that is iteratively refined using a measure of error in the predictions made by the model.Regression methods are a workhorse of statistics and have been co-opted into statistical machine learning. This may be confusing because we can use regression to refer to the class of problem and the class of algorithm. Really, regression is a process. The most popular regression algorithms are:

- Ordinary Least Squares Regression (OLSR)
- Linear Regression
- Logistic Regression
- Stepwise Regression
- Multivariate Adaptive Regression Splines (MARS)
- Locally Estimated Scatterplot Smoothing (LOESS)

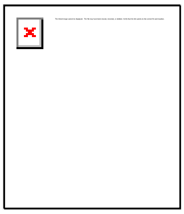
D. Instant-Based Algorithms



Instance-based learning model is a decision problem with instances or examples of training data that are deemed important or required to the model. Such methods typically build up a database of example data and compare new data to the database using a similarity measure in order to find the best match and make a prediction. For this reason, instance-based methods are also called winner-take-all methods and memory-based learning. Focus is put on the representation of the stored instances and similarity measures used between instances. The most popular instance-based algorithms are:k-Nearest Neighbor (kNN)

- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Locally Weighted Learning (LWL)
- Support Vector Machines (SVM)

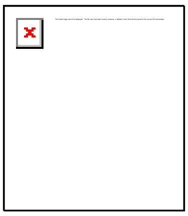
E. Regularization Algorithms



An extension made to another method (typically regression methods) that penalizes models based on their complexity, favoring simpler models that are also better at generalizing. I have listed regularization algorithms separately here because they are popular, powerful and generally simple modifications made to other methods. The most popular regularization algorithms are:Ridge Regression

- Least Absolute Shrinkage and Selection Operator (LASSO)
- Elastic Net
- Least-Angle Regression (LARS)

F. Decision Tree Algorithms

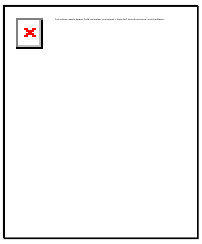


Decision tree methods construct a model of decisions made based on actual values

of attributes in the data. Decisions fork in tree structures until a prediction decision is made for a given record. Decision trees are trained on data for classification and regression problems. Decision trees are often fast and accurate and a big favorite in machine learning. The most popular decision tree algorithms are:
Classification and Regression Tree (CART)

- Iterative Dichotomiser 3 (ID3)
- C4.5 and C5.0 (different versions of a powerful approach)
- Chi-squared Automatic Interaction Detection (CHAID)
- Decision Stump
- M5
- Conditional Decision Trees

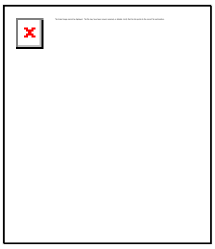
G. Artificial Neural Network Algorithms



Artificial Neural Networks are models that are inspired by the structure and/or function of biological neural networks.They are a class of pattern matching that are commonly used for regression and classification problems but are really an enormous subfield comprised of hundreds of algorithms and variations for all manner of problem types. Note that I have separated out Deep Learning from neural networks because of the massive growth and popularity in the field. Here we are concerned with the more classical methods. The most popular artificial neural network algorithms are:

- Perceptron
- Multilayer Perceptrons (MLP)
- Back-Propagation
- Stochastic Gradient Descent
- Hopfield Network
- Radial Basis Function Network (RBFN)

H. Deep Learning Algorithms



Deep Learning methods are a modern update to Artificial Neural Networks that exploit abundant cheap computation. They are concerned with building much larger and more complex neural networks and, as commented on above, many methods are concerned with very large datasets of labelled analog data, such as image, text. audio, and video. The most popular deep learning algorithms are:

- Convolutional Neural Network (CNN)
- Recurrent Neural Networks (RNNs)
- Long Short-Term Memory Networks (LSTMs)
- Stacked Auto-Encoders
- Deep Boltzmann Machine (DBM)
- Deep Belief Networks (DBN)

III. CONCLUSIONS

This paper proposes a context-aware network model based on transfer learning, which aims to predict learners' performance by solving their problems and improving the educational process, contributing to a comprehensive analysis of such student behavior and exploring various learning models. The scheme in this paper achieves a certain effect of personalized learning strategy, but the model is too large, and the model structure can be optimized in the future.

REFERENCES

[1] K. S. Hone and G. R. El Said, "Exploring the factors affecting MOOC retention: a survey study," *Computers & Education*, vol. 98, pp. 157–168, 2016.

[2] C. M. Stracke and A. Bozkurt, "Evolution of MOOC designs, providers and learners and the related MOOC research and publications from 2008 to 2018," in *Proceedings of the International Open & Distance Learning Conference (IODL19)*, Eskisehir, Turkey, January 2019.

[3] P. V. Murugan and V. M. Queen, "MOOCs as a digital learning platform," *International Journal of Multidisciplinary Research Review*, vol. 3, no. 1, pp. 28–37, 2020.

[4] T. Zhou, S. Huang, J. Cheng, and Y. Xiao, "The distance teaching practice of combined mode of massive open online course micro-video for interns in emergency department during the COVID-19 epidemic period," *Telemedicine and e-Health*, vol. 26, no. 5, pp. 584– 588, 2020.

[5] A. Moattarian, M. H. Tahririan, and A. Alibabaee, "Task complexity manipulation and EFL learners' interactions in the process of collaborative pre-planning," *Applied Research on English Language*, vol. 8, no. 1, pp. 51–78, 2019.

[6] L. Y. Li, "Effect of prior knowledge on attitudes, behavior, and learning performance in video lecture viewing," *International Journal of Human-Computer Interaction*, vol. 35, no. 4-5, pp. 415–426, 2019.

[7] J. Wong, M. Khalil, M. Baars, B. B. de Koning, and F. Paas, "Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course," *Computers & Education*, vol. 140, Article ID 103595, 2019.

[8] D. J. Lemay and T. Doleck, “Grade prediction of weekly assignments in MOOCS: mining video-viewing behavior,” *Education and Information Technologies*, vol. 25, no. 2, pp. 1333–1342, 2020.

[9] K. Limsathitwong, K. Tiwatthanont, and T. Yatsungnoen, “Dropout prediction system to reduce discontinue study rate of information technology students,” in *Proceedings of the 2018 5th International Conference on Business and Industrial Research (ICBIR)*, pp. 110–114, IEEE, Bangkok, Thailand, May 2018.

[10] V. Hegde and P. P. Prageeth, “Higher education student dropout prediction and analysis through educational data mining,” in *Proceedings of the 2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pp. 694–699, IEEE, Coimbatore, India, June 2018.

[11] G. Nagarajan and R. I. Minu, “Multimodal fuzzy ontology creation and knowledge information retrieval,” in *Proceedings of the International Conference on Soft Computing Systems*, pp. 697–706, Springer, Dubai, UAE, November 2016.

[12] T. Soffer and A. Cohen, “Students' engagement characteristics predict success and completion of online courses,” *Journal of Computer Assisted Learning*, vol. 35, no. 3, pp. 378–389, 2019.

[13] D. Ifenthaler and J. Y.-K. Yau, “Utilising learning analytics to support study success in higher education: a systematic review,” *Educational Technology Research & Development*, vol. 68, no. 4, pp. 1961–1990, 2020.

[14] M. Liz-Domínguez, M. Caeiro-Rodríguez, M. Llamas-Nistal, and F. A. Mikic-Fonte, “Systematic literature review of predictive analysis tools in higher education,” *Applied Sciences*, vol. 9, no. 24, 5569 pages, 2019.

[15] B. Li and X. Xu, “Application of artificial intelligence in basketball sport,” *Journal of Education, Health and Sport*, vol. 11, no. 7, pp. 54–67, 2021.

[16] R. A. Minhas, A. Javed, A. Irtaza, M. T. Mahmood, and Y. B. Joo, “Shot classification of field sports videos using AlexNet convolutional neural network,” *Applied Sciences*, vol. 9, no. 3, 483 pages, 2019.

[17] K. S. McCarthy, M. Watanabe, J. Dai, and D. S. McNamara, “Personalized learning in istart: past modifications and future design,” *Journal of Research on Technology in Education*, vol. 52, no. 3, pp. 301–321, 2020.

[18] J. D. Basham, T. E. Hall, R. A. Carter, and W. M. Stahl, “An operationalized understanding of personalized learning,” *Journal of Special Education Technology*, vol. 31, no. 3, pp. 126–136, 2016.

[19] P. A. Bishop, J. M. Downes, S. Netcoh et al., “Teacher roles in personalized learning environments,” *The Elementary School Journal*, vol. 121, no. 2, pp. 311–336, 2020.

[20] S. F. Tseng, Y. W. Tsao, L. C. Yu, C. L. Chan, and K. R. Lai, “Who will pass? analyzing learner behaviors in MOOCs,” *Research and Practice in Technology Enhanced Learning*, vol. 11, no. 1, pp. 1–11, 2016.

[21] R. Al-Shabandar, A. J. Hussain, P. Liatsis, and R. Keight, “Analyzing learners behavior in MOOCs: an examination of performance and motivation using a data-driven approach,” *IEEE Access*, vol. 6, pp. 73669–73685, 2018.

[22] M. L. Bote-Lorenzo and E. Gómez-Sánchez, “Predicting the decrease of engagement indicators in a MOOC,” in *Proceedings of the 7th International Learning Analytics & Knowledge Conference*, pp. 143–147, Vancouver, Canada, March 2017.

[23] C. Zhang, T. Xie, K. Yang et al., “Positioning optimisation based on particle quality prediction in wireless sensor networks,” *IET Networks*, vol. 8, no. 2, pp. 107–113, 2019.

[24] Z. Zhang, C. Zhang, M. Li, and T. Xie, “Target positioning based on particle centroid drift in large-scale WSNs,” *IEEE Access*, vol. 8, pp. 127709–127719, 2020.

[25] C. D. Nguyen and T. C. T. Dang, “Second language teacher education in response to local needs: preservice teachers of English learning to teach diverse learners in communities,” *Tesol Quarterly*, vol. 54, no. 2, pp. 404–435, 2020.

[26] Y. Madani, M. Erritali, J. Bengourram, and F. Sailhan, “Social collaborative filtering approach for recommending courses in an E-learning platform,” *Procedia Computer Science*, vol. 151, pp. 1164–1169, 2019.

[27] N. A. Albatayneh, K. I. Ghauth, and F. F. Chua, “Utilizing learners’ negative ratings in semantic content-based recommender system for e-learning forum,” *Journal of Educational Technology & Society*, vol. 21, no. 1, pp. 112–125, 2018.

[28] L. Chen and W. Wang, “Analysis of technical features in basketball video based on deep learning algorithm,” *Signal Processing: Image Communication*, vol. 83, Article ID 115786, 2020.

[29] A. Tlili, M. Denden, F. Essalmi et al., “Does providing a personalized educational game based on personality matter? A case study,” *IEEE Access*, vol. 7, pp. 119566–119575, 2019.

[30] Y. Fan, X. Lu, D. Li, and Y Liu, "Video-based emotion recognition using CNN-RNN and C3D hybrid networks," in *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, pp. 445–450, Tokyo Japan, November 2016.

[31] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, no. 1, pp. 1–40, 2016.

[32] A. Cohen, "Analysis of student activity in web-supported courses as a tool for predicting dropout," *Educational Technology Research & Development*, vol. 65, no. 5, pp. 1285–1304, 2017.