INDO ASIAN JOURNAL OF INFORMATION RESEARCH AND TECHNOLOGY

VOL 1 (01), AUG 2024, PP. 4-13

SUPERVISED LEARNING ALGORITHMS: AN OVERVIEW

UPASANA GUPTA
EDITORIAL OFFICE, BALLPEN MEDIA PVT LTD
EDITORIALOFFICE@BALLPENMEDIAPVTLTD.CO.IN

ABSTRACT

In the realm of machine learning, numerous supervised learning techniques have been established in the past ten years. A lot of machine learning research is now being conducted in the field of supervised learning. Numerous supervised learning approaches have found use in the processing and analysis of several kinds of data. The capacity of supervised learning to use annotated training data is one of its key features. In the categorisation process, the so-called labels are class labels. The supervised learning techniques employ a wide range of algorithms. In this study, the basic ideas behind a few supervised approaches are outlined. This review paper's primary objective is to give an overview of machine learning technique.

Keywords: Machine Learning, classification, Supervised techniques

1. INTRODUCTION

Information technology, statistics, probability, artificial intelligence, psychology, neuroscience, and many more fields all present a sizable area of study called machine learning. By creating a model that accurately represents a chosen dataset, machine learning can address problems with ease(Khan et al, 2010). From training computers to simulate the human brain, machine learning has developed into a cutting-edge field that has elevated statistics to a comprehensive discipline that generates basic statistical computational theories of learning processes.

Developing methods that let computers learn is the core of machine learning. Finding statistical regularities or other patterns in data is the process of learning. The purpose of machine learning algorithms(Talwar and Kumar, 2013) is to simulate how a human might go about learning a particular activity. These algorithms can also provide insight into how challenging learning can be in certain settings.

These days, machine learning is different from what it used to be due to the advancement of new computing technologies in the field of big data. A recent advancement in machine learning is the capacity to automatically apply a range of complex mathematical calculations to a vast data, which calculates the results more faster. Many machine learning algorithms have been invented, updated, and improved (Caruana et al, 2006).

Adaptive programming is incredibly well-liked. It is applied to machine learning applications that can identify patterns, learn from past mistakes, extract new knowledge from data, or maximize the precision and effectiveness of data processing and output. Additionally, a variety of application domains use multidimensional data, which is worked with through machine learning techniques (Ayodele, 2010).

Thus, the machine learning algorithms are grouped into the following categories according to the intended result of the algorithm:

Supervised learning: The different methods provide a function that maps inputs to desired outputs, which is known as supervised learning. The classification issue is one common way to formulate a supervised learning task: the learner must study multiple input-output instances of a function in order to learn (or approximate the behaviour of) the function that maps a vector into one of several classes.

Unsupervised learning: Unsupervised learning involves modelling a set of inputs in the absence of labelled examples.

Semi-supervised learning: This method creates a suitable function or classifier by combining instances that have been tagged and those that have not.

In supervised algorithms, the classes are predetermined. These classes are created in a manner of finite set, defined by the human, which in practice means that a certain segment of data will be labeled with these classifications. The task of the machine learning algorithm is to find patterns and construct mathematical models. These models are then evaluated based on the predictive capacity in relation to measures of variance in the data itself. It is also useful to make difference between two main supervised models: classification models (classifiers) and regression models. Regression models map the input space into a real-value domain. The classifiers map the input space into pre-defined classes. There are many alternatives for representing classifiers, for instance, support vector machines, decision trees, probabilistic summaries, algebraic function, etc. Along with regression and probability estimation, classification is one of the most studied models, possibly one with the greatest practical relevance. The potential benefits of progress in classification are immense since the technique has great impact on other areas, both within Data Mining and in its applications(Dhage and Raina, 2016). On the other hand, the unsupervised learning algorithms are not provided with classifications. The main task of unsupervised learning is to automatically develop classifications labels. These algorithms are searching for the similarity between pieces of data in order to determinate if they can be categorized and create a group.

These groups are so called clusters, and they represent whole family of clustering machine learning techniques. In this unsupervised classification (cluster analysis) the machine doesn't know how the clusters are grouped. Using the cluster analysis, there is a bigger potential for surprising ourselves. Thus, cluster analysis is a very promising tool for the exploration of relationships between many papers.

2. MATERIALS AND METHODS

Table I

Research Title	Author(s)	Year	Citation	Reference	Similarity to original
Supervised Learning in Physical Networks: From Machine Learning to Learning Machines	M. Stern, D. Hexner, J. Rocks, Andrea J. Liu	2020	60	49	100
Demonstration of Decentralized Physics-Driven Learning	Sam Dillavou, M. Stern, Andrea J. Liu, D. Durian	2021	42	65	49
Directed aging, memory, and nature's greed	Nidhi Pashine, D. Hexner, Andrea J. Liu, S. Nagel	2019	72	43	30.6
Supervised learning through physical changes in a mechanical system	M. Stern, Chukwunonso Arinze, Leron Perez, S. Palmer, A. Murugan	2020	49	48	30.3

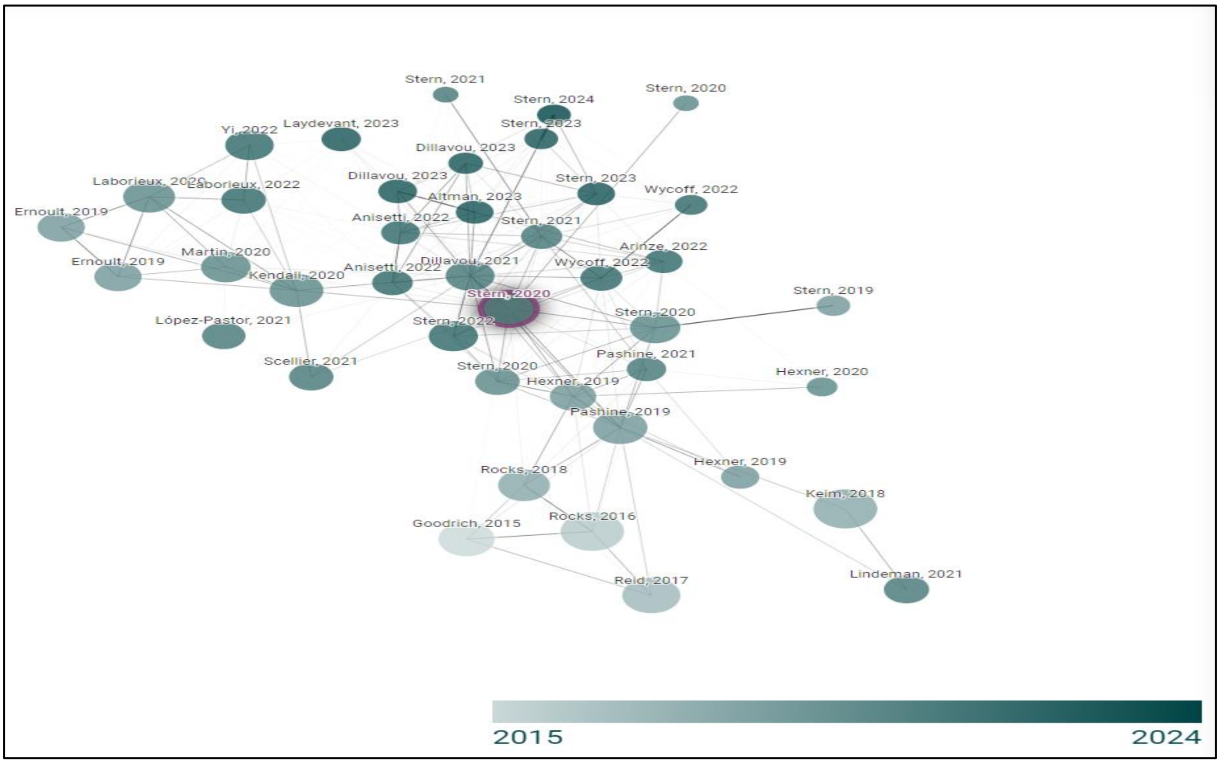
Learning Without Neurons in Physical Systems	M. Stern, A. Murugan	2022	43	134	29.3
Periodic training of creeping solids	D. Hexner, Andrea J. Liu, S. Nagel	2019	31	50	27.4
Training End-to-End Analog Neural Networks with Equilibrium Propagation	Jack D. Kendall, Ross D. Pantone, Kalpana Manickavasagam, Yoshua Bengio, B. Scellier	2020	69	56	27
Flow networks as learning machines	M. Stern, D. Hexner, J. Rocks, Andrea J. Liu	2020	0	14	26.3
Physical learning beyond the quasistatic limit	M. Stern, Sam Dillavou, Marc Z. Miskin, D. Durian, Andrea J. Liu	2021	15	28	25.6
Continual Learning of Multiple Memories in Mechanical Networks	M. Stern, Matthew B. Pinson, A. Murugan	2020	25	45	25
Supervised learning in a mechanical system	M. Stern, Chukwunonso Arinze, Leron Perez, S. Palmer, A. Murugan	2019	3	41	23.4
Desynchronous learning in a physics-driven learning network.	J. F. Wycoff, Sam Dillavou, M. Stern, A. Liu, D. Durian	2022	17	44	23.1
Experimental Demonstration of Coupled Learning in Elastic Networks	Lauren E. Altman, M. Stern, Andrea J. Liu, D. Durian	2023	8	49	22.9
Machine Learning Without a Processor: Emergent Learning in a Nonlinear Electronic Metamaterial	Sam Dillavou, Benjamin D Beyer, M. Stern, Marc Z. Miskin, Andrea J. Liu, D. Durian	2023	10	47	21.2
A deep learning theory for neural networks grounded in physics	B. Scellier	2021	24	0	21.1
Learning by non-interfering feedback chemical signaling in physical networks	Vidyesh Rao Anisetti, B. Scellier, J. Schwarz	2022	14	61	21
Circuits that train themselves: decentralized, physics-driven learning	Sam Dillavou, Benjamin D Beyer, M. Stern, Marc Z.	2023	4	15	20.6

	Miskin, A. J. Liu, D. Durian				
Physical learning of power-efficient solutions	M. Stern, Sam Dillavou, Dinesh Jayaraman, D. Durian, Andrea J. Liu	2023	3	57	20.3
Training self-learning circuits for power-efficient solutions	M. Stern, Sam Dillavou, Dinesh Jayaraman, D. Durian, Andrea J. Liu	2024	3	58	20.2
Local rules for fabricating allosteric networks	Nidhi Pashine	2021	11	18	19.8
The Physical Effects of Learning	M. Stern, A. J. Liu, V. Balasubramanian	2023	8	76	19.4
Effect of aging on the non-linear elasticity and memory formation in materials	D. Hexner, Nidhi Pashine, Andrea J. Liu, S. Nagel	2019	9	47	18.5
Learning Without a Global Clock: Asynchronous Learning in a Physics-Driven Learning Network	J. F. Wycoff, Sam Dillavou, M. Stern, Andrea J. Liu, D. Durian	2022	2	36	17.8
EqSpike: spike-driven equilibrium propagation for neuromorphic implementations	Erwann Martin, M. Ernoult, Jérémie Laydevant, Shuai-shuai Li, D. Querlioz, Teodora Petrisor, J. Grollier	2020	45	65	17.6
Limits of multifunctionality in tunable networks	J. Rocks, Henrik Ronellenfitsch, Andrea J. Liu, S. Nagel, E. Katifori	2018	57	54	17.1
Out of equilibrium learning dynamics in physical allosteric resistor networks	M. Stern	2021	0	25	16.9
Scaling Equilibrium Propagation to Deep ConvNets by Drastically Reducing Its Gradient Estimator Bias	Axel Laborieux, M. Ernoult, B. Scellier, Yoshua Bengio, J. Grollier, D. Querlioz	2020	57	42	16.8
Learning to self-fold at a bifurcation.	Chukwunonso Arinze, M. Stern, S. Nagel, A. Murugan	2022	9	56	16.5
Frequency Propagation: Multimechanism Learning in	Vidyesh Rao Anisetti, A.	2022	10	37	16.1

Nonlinear Physical Networks	Kandala, B. Scellier, J. Schwarz				
Updates of Equilibrium Prop Match Gradients of Backprop Through Time in an RNN with Static Input	M. Ernoult, J. Grollier, D. Querlioz, Yoshua Bengio, B. Scellier	2019	35	22	15.9
Equilibrium Propagation with Continual Weight Updates	M. Ernoult, J. Grollier, D. Querlioz, Yoshua Bengio, B. Scellier	2019	35	20	15.5
Training nonlinear elastic functions: nonmonotonic, sequence dependent and bifurcating.	D. Hexner	2020	1	40	15.2
Memory formation in matter	N. Keim, Joseph D. Paulsen, Z. Zeravcic, S. Sastry, S. Nagel	2018	155	176	15.1
Training an Ising machine with equilibrium propagation	Jérémie Laydevant, Danijela Marković, J. Grollier	2023	11	79	14.8
Auxetic metamaterials from disordered networks	Daniel R. Reid, Nidhi Pashine, J. Wozniak, H. Jaeger, Andrea J. Liu, S. Nagel, J. D. de Pablo	2017	99	34	14.6
Multiple memory formation in glassy landscapes	C. Lindeman, S. Nagel	2021	27	31	14.5
Designing allostery-inspired response in mechanical networks	J. Rocks, Nidhi Pashine, Irmgard Bischofberger, C. Goodrich, Andrea J. Liu, S. Nagel	2016	150	33	14.4
Self-learning Machines based on Hamiltonian Echo Backpropagation	V. López-Pastor, F. Marquardt	2021	21	108	14.2
Holomorphic Equilibrium Propagation Computes Exact Gradients Through Finite Size Oscillations	Axel Laborieux, F T Zenke	2022	25	85	14.2
Activity-difference training of deep neural networks using memristor crossbars	Su-in Yi, Jack D. Kendall, R. S. Williams, Suhas Kumar	2022	40	58	13.9
The Principle of Independent	C. Goodrich,	2015	84	4	13.7

Bond-Level Response: Tuning by Pruning to Exploit Disorder for Global Behavior.	Andrea J. Liu, S. Nagel				

Figure I Scietific literature related to supervided leraning



Source: Author Compilation

3. DISCUSSION

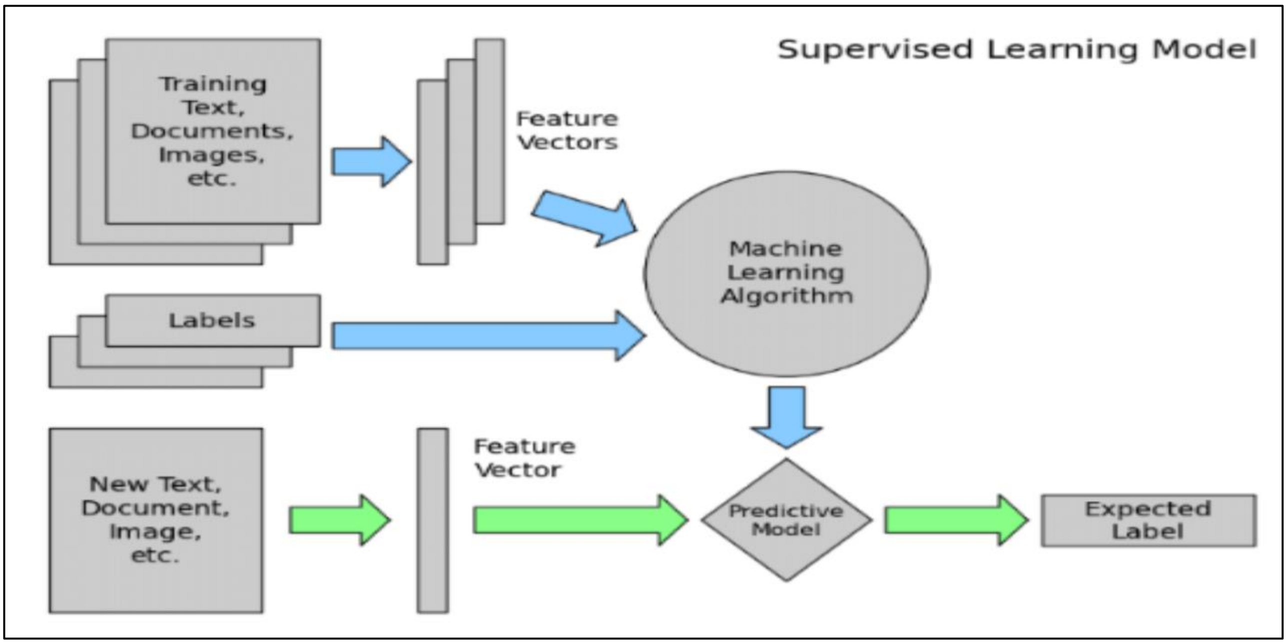
3.1 How supervised learning Work

The learning process in a simple machine learning model is divided into two steps: training and testing. In training process, samples in training data are taken as input in which features are learned by learning algorithm or learner and build the learning model(Maimon and Rokach, 2010). In the testing process, learning model uses the execution engine to make the prediction for the test or production data. Tagged data is the output of learning model which gives the final prediction or classified data. Supervised learning (Figure 1) is the most common technique in the classification problems, since the goal is often to get the machine to learn a classification system that we've created. Most commonly, supervised learning leaves the probability for input undefined, such as an input where the expected output is known. This process provides dataset consisting of features and labels(Ayodele, 2010). The main task is to construct an estimator able to predict the label of an object given by the set of features. Then, the learning algorithm receives a set of features as inputs along with the correct outputs and it learns by comparing its actual output with corrected outputs to find errors. It then modifies the model accordingly. The model that is created is not needed as long as the inputs are available, but if some of the input values are missing, it is not possible to infer anything about the outputs. Supervised learning is the most common technique for training for neutral networks and decision trees(Akritidis and Bozanis, 2013). Both of these are depended on the information given by the pre-determinate classification. Also, this learning is used in applications where historical data predicts likely feature events(Rokach and Maimon, 2008). There are many practical examples of this learning, for instance an application that predicts the species of iris given a set of measurements of its flower.

As previously mentioned, the supervised learning tasks are divided into two categories: classification and regression. In classification, the label is discrete, while in regression, the label is continuous(Caruana, 2006). As shown on Figure 2, the algorithm makes the distinction between the observed data X that is the training data, in most cases structured data given to the model during the training process. In this process, the supervised learning algorithm builds the predictive model. After its training, the fitted model would try to predict the most likely labels for a new set of samples X in the testing set. Depending on the nature of the target y, supervised learning can be classified:

- If y has values in a fixed set of categorical outcomes (integers), the task to predict y is called classification
- If y has floating point values, the task to predict y is called regression

Figure II Supervised Learning Model



Source: (Kotsiantis et al, 2007)

3.2 Application of supervised learning

There are many different issues that can be resolved by supervised learning, such as:

- Spam filtering: By training supervised learning algorithms to recognize and categorize spam emails according to their content, users can steer clear of unsolicited communications.
- Supervised learning makes picture classification easier by automatically classifying images into categories like objects, animals, or scenes. This makes activities like image search, content moderation, and image-based product recommendations easier to do.
- Medical diagnosis can be aided by supervised learning, which analyzes patient data, including test results, medical pictures, and patient histories, to find patterns that point to particular diseases or disorders.
- Fraud detection: Financial institutions can safeguard their clients and stop fraud by using supervised learning models to examine financial transactions and spot trends that point to fraudulent conduct.
- Natural language processing (NLP): In order for robots to successfully comprehend and interpret human language, supervised learning is essential for NLP tasks including sentiment analysis, machine translation, and text summarization.

4. CONCLUSION

According to the paper's discussion, supervised learning is one of the most often used machine learning methodologies. The annotated training data makes it possible to establish better criteria for model optimization, which makes the employed procedures much more effective than unsupervised ones. Many algorithms are used in supervised learning techniques, and data scientists are constantly refining these algorithms. The machine learning procedure is briefly explained. Many developers are paying close attention to this sector, which has advanced significantly over the past ten years. Excellent results were attained by the learning strategies, which would not have been possible in earlier decades. There is lots of room for the developers to enhance or refine the algorithms and supervised learning methods because of the quick progress.

REFERENCES

Ayodele, T. O. (2010). Types of machine learning algorithms in new advances in machine learning. croatia, rijeka.

Akritidis, L., & Bozanis, P. (2013, March). A supervised machine learning classification algorithm for research articles. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing* (pp. 115-120).

Caruana, R., & Niculescu-Mizil, A. (2006, June). An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on Machine learning* (pp. 161-168).

Khan, A., Baharudin, B., Lee, L. H., & Khan, K. (2010). A review of machine learning algorithms for text-documents classification. *Journal of advances in information technology*, *1*(1), 4-20.

Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, *160*(1), 3-24.

Maimon, O. Z., & Rokach, L. (2014). *Data mining with decision trees: theory and applications* (Vol. 81). World scientific.

Sandhya, N., & Charanjeet, K. R. (2016). A review on machine learning techniques. *International Journal on Recent and Innovation Trends in Computing and Communication*, *4*(3), 451-458.

Talwar, A., & Kumar, Y. (2013). Machine Learning: An artificial intelligence methodology. *International Journal of Engineering and Computer Science*, *2*(12), 3400-3404.