A ADVANCED JOURNAL IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING TRENDS IN BIOLOGICAL SCIENCES
Open Journal

Review

An Investigation into the Techniques and Uses of Machine Learning

Serthan S. Altunaydin*

Osmangazi University, Eskisehir, Turkey

Corresponding author
Serthan S. Altunaydin
Osmangazi University, Eskisehir, Turkey

Article information
Received: September 20th, 2022; **Revised:** December 15th, 2022; **Accepted:** January 1st, 2023; **Published:** February 18th, 2023

Cite this article
Altunaydin SS. An investigation into the techniques and uses of machine learning. 2023; 2(1). doi: https://doi.org/10.70705/ppp.bioai.2023.v02.i01.pp8-13

ABSTRACT

A branch of AI, machine learning emerged in the 1950s and has since grown into its own scientific discipline. Machine learning's origins may be traced to the 1950s, when the field saw little in the way of substantial study or advancements. Nevertheless, studies in this area were revived, advanced, and continue to this day in the 1990s. This field of study has a lot of room for growth in the years to come. The difficulty in analyzing and interpreting the ever-increasing data is the driving force behind this change. Increasing amounts of data allow machine learning to operate on the premise of selecting the most appropriate model from among existing data in order to apply to new data. Consequently, studies in machine learning will continue in tandem with the growth of data. Machine learning's background, its methodologies, its application domains, and previous studies are all part of this body of work. Learning about machine learning and its applications has recently grown in popularity, and this project aims to teach academics about it.

Keywords
Machine learning; Machine learning algorithms; Artificial intelligence; Big data.

INTRODUCTION

T"he process of a change and enhancement in the behaviours through exploring new information in time" is how Simon defines learning. The term "machine learning" is used when the "learning" in this definition is carried out by computers. During the machine learning process, the word "enhancement" refers to developing the optimal answer by using past experiences and samples (Sırmaçek, 2007). The phrase "big data" came into use because of advancements in IT. According to Altunışık (2015), the idea of "big data" is not novel; it refers to massive, ever-growing datasets that defy analysis using conventional database methods. The data acquired from various online apps, automated teller machines, credit card scanners, and so on is massive. Analysis of the data gathered in this manner is imminent. The goals of analyzing data from various sectors vary across different industries. Natural language processing, image processing, computer vision, voice and handwriting recognition, manufacturing, energy generation, calculating finance, biology, and the automobile, aviation, and industrial industries are just a few of the many that make use of machine learning applications. Nonetheless, the goal is predicated on the idea of deducing meaning from prior facts. Machine learning algorithms and techniques were created to analyze and understand data since humans are unable to do so (Amasyalı, 2008).

The idea of machine learning, which has been all the rage as of late, is dissected here. Machine learning's background, methodology, and algorithms, as well as its many applications, are covered in the study. The last section is a conclusion that summarizes the findings from all the prior research.

2. MACHINE LEARNING

2.1. Definition

There is no error margin in the operations carried out by computers based an algorithm and the operation follows certain steps. Different from the commands which are written to have an output based on an input, there are some situations when the computers make decisions based upon the present sample data. In those situations, computers may make mistakes just like people in the decision-making process. That is, machine learning is the process of equipping the computers with the ability to learn by using the data and experience like a human brain (Gör, 2014).

The main aim of machine learning is to create models which can train themselves to improve, perceive the complex patterns, and find solutions to the new problems by using the previous data (Tantuğ ve Türkmenoğlu, 2015).

2.2. History

In 1940s, based on the studies on the electrical crashes of the neurons, the scientists explained the decision-making mechanism of human by cannon and fire. In this way, the researches of the artificial intelligence started in the 1950s (Erdem, 2014). In those years, Alan Turin executed the Turing Test in order to test the ability of a machine to imitate a human. The aim of the Turing Test was to measure the ability of the machine to make a contact with a human during an interview. If the machine performed worse than a human, it was successful. In 1956, the term 'artificial intelligence' was first used in a summer school held by Marvin Minsky from Massachusetts Institute of Technology, John McCarthy from Stanford University and Allen Newell and Herbert Simon from Carnegie-Mellon University. Until that time, Alan Turing's term, 'machine intelligence', had been used. In 1959, Arthur Samul created the checkers programme, and then machine learning got its way. From those developments to the 1980s, there were some studies on abstract mind, information-based systems, which was called the 'winter of artificial intelligence'. In the 1990s, artificial intelligence and machine learning studies accelerated due to the developments in game technologies. Nowadays, artificial intelligence and machine learning are used in lots of researches and work sectors (Topal, 2017)

2.3. Machine Learning Methods

Machine Learning can be examined in four parts as follows;

- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Reinforced learning

Supervised Learning: It is a method in which the present input data is used to reach the result set. There are two types of supervised learning: classification and regression supervised learning.

- Classification: Distributing the data into the categories defined on the data set according to their specific features.
- Regression: Predicting or concluding the other features of the data based on its some available features.

Unsupervised Learning: The difference between the supervised and unsupervised learning is that in unsupervised learning the output data is not given. The learning process occurs by using the relations and connections between the data. Also, unsupervised learning doesn't have a training data.

There are also two types of unsupervised learning: clustering and association.

- Clustering: Finding the groupings of data which are similar to each other when inherent groupings in the data is not known.
- Association: Determining the relations and connections among the data in the same data set.

Deduction of Features: In some cases, although lots of features about the data are known, the features related to group and category of the data cannot be determined. In such cases, selecting a subgroup of features or getting new features combining the features is called deduction of features (Erdem, 2014).

Semi-supervised Learning: supervised and unsupervised learning is inadequate when the labelled data are less than unlabelled data. In such cases, the unlabelled data, which are very inadequate, is used to deduce information about them. And, this method is called semi-supervised learning. The difference between the semi-supervised learning and the supervised learning is the labelled data set. In supervised learning, the labelled data are more than the data to be predicted. In contrast, in semi-supervised learning, the labelled data are less than the data to be predicted (Kızılkaya ve Oğuzlar, 2018).

Reinforcement Learning: This is a kind of learning in which the agents learn via reward system. Although there is a start and finish points, the aim of the agent is to use the shortest and the correct ways to reach the goal. When the agent goes through the correct ways, s/he is given positive rewards. But the going through wrong ways means negative rewards. Learning occurs on the way to the goal (Sırmaçek, 2007).

2.4. Machine Learning Algorithms

2.4.1. Artificial Neural Networks

Artificial neural network is a data processing system which is developed based on the biological neural networks in the human brain to function like human brain neural networks (Kocadayı, Erkaymaz, ve Uzun, 2017).

Neurons (process elements) are the basics of artificial neural networks. Neurons have 5 basic functions: inputs, weights, summation function, activation function and output.

Activation Function: This function is used to calculate the output value which corresponds the input value. In some neural network models, it is must for the activation function to be derivable. Calculating the derivative is important for the learning process of the network. Thus, the derivation of the sigmoid function is the most commonly used function because it can be written in the function itself. It is not compulsory to use the same activation function in all the cells. They can have different activation functions. Activation functions are as follows: linear function, sigmoid function, hyperbolic tangent function, sine function, digit function.

Output: This is the value which is determined by the activation value. The last output produced can both be sent to the other cells or to the outer world. If there is a feedback, the cell may use the output as an input by this feedback (Haciefendioğlu, 2012).

2.4.1.1. Single Layer and Multilayer Artificial Neural Networks

The first researches on artificial intelligence started with the single layer artificial neural networks. The most important feature of the network is classification of the problems which can be selected linear as a layer. After the inputs in the problem are multiplied by the weights and added, the calculated values are classified according to their threshold value as high or low. The groups are shown like -1 and 1 or 0 and 1. During the learning process, both the weights and the weights of threshold value are updated. The output value of the threshold value is 1. Since the single layer artificial neural networks are inefficient for the nonlinear problems, multilayer artificial neural networks have been developed. Today, mostly used artificial neural network is the multilayer artificial neural network. Multilayer networks emerged during the studies to solve the XOR problems. Multilayer networks have 3 layers.

Input Layer: This layer gets the information from the outer world, but there is no process on this layer.

Interlayers: The information from the input layer is processed on this layer. Mostly one interlayer can be adequate for the solution of the problem. However, if the relations between input and output are not linear or there are some complications, more than one layer can be used.

Output Layer: The information from the interlayer is processed on this layer and the outputs which correspond the input are detected.

In training the multilayer artificial neural networks, the 'delta rule' is used. As the multilayer networks use supervised learning methods, both the inputs and the outputs which correspond the inputs are shown to the network. According to the learning rule, the error margin between the outputs and the expected outputs are distributed to the network in order to minimize the error margin (Öztemel, (2003).

2.4.1.2. Feedforward and back propagation artificial neural networks

Artificial neural network architectures are divided into two groups as feedforward and back propagation based on the directions of the links between the neurons. In the feedforward networks, the signals go from input layer to output layer on the one-way links. At the same time, in the feedforward networks, the output values of the cells in one layer are transmitted to the following layers as the inputs on the weights. The input layer sends the input to the hidden layer without making any change. Once this information is processed on the hidden and the output layer, its output on the network is determined. Multilayer sensors and learning vector quantity can be examples of feedforward artificial networks.

The most important characteristics of the back propagation artificial neural networks is that output value of at least one cell is given to itself or another cell as an input value. The back propagation can be processed on a retardation unit as well as the cells in one layer or among the cells on other layers. Because of this feature, the back propagation artificial neural networks show a dynamic behaviour [12]. Those networks got their name by their function that they can organize the weights backwards in order to minimize the errors occurred on the output layer (Hamzaçebi ve Kutay, 2004).

2.4.2. Decision Trees

A decision tree which learns from the data classified by the induction is a decision making structure. It is a kind of learning algorithm which divides the large amount of data into small portions by using simple decision making steps. At the end of every successful division, the similarity of the elements in the final group increases. The decision trees, which have descriptive and predictive features, are one of the most popular classification algorithms because they can be easily interpreted, integrated to the databases and are reliable (Albayrak ve Yılmaz Koltan, 2009). Decision tree have three structures: decision nodes, branches and leaves.

Root Nodes: It is a node which has no former branch and can create one or more branch. Root nodes show the dependent variable and show which variable will be used for the classification.

Interior Node: It is a node which has one incoming branch and can have two or more outgoing branches.

Leaf or Terminal nodes: These are the nodes which has an incoming branch but no outgoing brand.

This is a structure which shows the result of the test between the leaves and the nodes, and has a role to determine the groups to be defined. If the classification is not completed at the end of the branch, a decision node emerges. The place of the nodes at the end of every branch is called deepness. The user can determine the number of deepness by analysing the appropriateness of the decision tree to the data set. In the decision trees, the deepth and the number of groups are directly proportional.

The decision tree is shaped by the questions and their answers. As a result, some rules emerge according to the answers. Once the variable, which is the source of the question, is determined, this variable creates the root node of the tree. The test to be applied is determined by the root node. At the end of the test, the tree is divided into branches and the separation process follow the test. Each of the branches on the are candidates for classification process. If there is a classification at the end of a branch, a leaf emerges at the end of the branch. The leaf is the one of the desired groups in the data. If there is no classification process at the end of the branch, there emerges a decision node on this branch. The decision tree aims to reach the leaf by the shortest way starting from the root node through sequencing nodes.

Each feature is used as a test in order to decide on the classification of the training data. After the best feature is chosen, it is used on the root node for the test. The number of branches changes according to the value of the feature. Which feature is going to be chosen on each node is the main selection of the decision tree. The measure of the feature is determined by a value called information gain which is also defined as entropy.

Entropy: Measuring the disorder in a system or events is called entropy. Entropy is related to the information and when uncertainty and disorder rise, more information is needed to define the data better. The value of the entropy changes between 0 and 1, and the value near 1 means more uncertainty. Therefore, it is necessary to lower the entropy value to 0 in decision trees. When D represents the distribution of probability P (p1,p2,….pn), the entropy equation is as follows:

As it is seen in the equation, E(D) represents the entropy before the dataset is divided; i represents the entropy of the subdivision after it is divided E(Di); p(Di) is the probability of the i subdivision after it is divided.

Pruning: overfitting may occur when creating a model on the decision tree. While the model becomes successful for the sample data, it can make mistakes with the new data. It occurs when there is too much information to be classified or noisy data in the dataset. Pruning is the process of cutting the branches which are formed by the noisy data and which leads to mistakes. The pruning process has two types: pre-pruning and post-pruning. Generally, post- pruning is preferred. In this process, determined branches are cut or two different branches are combined and cut after the whole tree has been created till the leaves by using the whole data. At the end of the pruning process the tree gets smaller with less error margins (Haciefendioğlu, 2012).

Widely used various decision tree methods are given in the following table:

2.4.3. Support Vector Machines

Support vector machines (SVM) are one of the supervised classification techniques which were founded by Cortes and Rapnik in 1995. SVM is a kind of machine algorithm which makes predictions and generalizations on the new data by learning on the data sets whose distribution is unclear. The main principle of SVM is based on finding the hyperplane which separates the data of two classes the most appropriately. Support vector machines are divided into two categories based on the classification that the data set is separated linearly and not linearly (Güneren, 2015).

Linearly separable case: With SVM it is aimed to separate the samples of two classes which are generally shown with the labels (-1, +1) with two different most appropriate hyperplane by the help of decision function generated at the end of the training data. This process is reached by finding the hyperplane which makes the length between the nearest spots to the SVM maximum. The hyperplane, which makes the border maximum, optimum hyperplane and the spots limiting the border are called support vectors.

2.5. Machine Learning Application Areas

The previous section includes the theoretical background of the machine learning algorithms. In this section, information about the areas and studies in which the machine learning are used nowadays will be given. Today, the use of machine learning has increased considerably. Although it is though that it can only be done in large studies, many people face machine learning in their daily life. These studies and applications are as follows:

Education: One of the most important application fields is education in which there have been some studies in order to identify and increase success recently. Despite the projects made in the field of education in recent years, the desired success has not been achieved. There are a lot of factors that influence this failure. However, it has not been determined which factor has more influence on this failure. In this context, by a questionnaire applied to secondary school students, the successes of the students in the lessons were predicted by machine learning models, which resulted with success (Gök, 2017).

Similarly, there are some studies in order to determine the proficiencies of students in higher education. In 2007, a study was carried out at Pamukkale University, where the students identified as risky students according to the failure in mathematics course. In the study, it was found out that the scores of 434 students' university entrance exam; mathematics, sciences, Turkish tests and high school graduation scores played a major role in predicting the success in mathematics. In the study, 289 students' data were used for training and 145 students' data were used for testing. As a result, 86 percent of the students who passed the mathematics course were correctly estimated (Güner ve Çomak, 2011).

Other areas of application for machine learning which have become quite functional in the field of education are:

Image processing: In this method, it is aimed to process and improve recorded images. Some application areas where the image processor is used are as follows:

- Security systems
- Face detection
- Medicine (to diagnose diseased tissues and organs)
- Military (to process underwater and satellite images)
- Motion detection
- Object detection

Computational biology:

- DNA sequencing
- Finding a tumor
- Drug discovery

Natural language processing: It is aimed to investigate and analyse the structures of natural languages. It is possible to perform many applications with natural language processing:

- Automatic translation of written texts
- Question-answer machines
- Automatic summarization of text
- Understanding speech and command

Automotive, aviation and production:

- Detecting malfunctions before they occur
- Producing autonomous vehicles

Retail:

- Customized shelf analysis for persons
- Recommendation engines
- Material and stock estimates
- Purchasing - demand trends

Finance:

- Credit controls and risk assessments
- Algorithmic trading

Agriculture:

- Predicting yields or deficiencies by analysing satellite images

Human Resources:

- Selecting the most successful candidate among a lot of applicants.

Energy:

- Calculating the heating and cooling loads for building designs
- Power usage analysis
- Smart network managements

Meteorology:

- Weather forecast via sensors

Health:

- Providing warning and diagnosis by analysing patient data
- Disease defining
- Health care analysis

Cyber security:

- Detecting the harmful network traffic
- Finding out address fraud

3. CONCLUSION

Machines have played an increasingly important part in our daily lives, driven by technological advancements in the last several years. All aspects of our life generate massive amounts of data, and this data is only growing. All of this data is being put to good use because of the machines. Contrary to popular belief, these devices are ubiquitous and employed in every aspect of human existence, not only engineering and computer science. Today, successful businesses are actively using this technology since they have previously identified

and invested in it. Lots of industries and individuals will be impacted in the future by computers that are good at tasks that humans aren't good at. Some well-established industries will go their own way, while others may give rise to entirely new ones. Information technology and machine power must be carefully considered in this setting.

REFERENCES

Yılmaz Koltan and Albayrak (2009). Evidence-Based Practices: Using Various Algorithms and Methods to Verify Data. Articles published in the Süleyman Demirel University Research and Academic Journal, volume 14, issue 1, pages 31-52.
(R. Altunışık) (2015). Is Büyük Veri: Fırsatlar Kaynağı mı Yoksa Değişilir?.
Published in the Yildiz Social Science Review, volume 1, issue 1, pages 48.

"Yeni Makine Öğrenmesi Metotları ve İlaç Tasarımına Uygulamaları" (2008, Amasyalı, M. F.).

Thesis Holder, Yıldız Technical University, Istanbul.

Ataseven, B. (2013). Marmara University Açık Arşiv Sistemi, 10(39), 101-115. Yapay Sinir Ağları ile Öngörü Modellemesi.

Published in 2013 by Çalış, K., Gazdağı, O., and Yıldız, O. Automatic Tests for Commercial Emails and Their Metin Madenciliği Yöntemleri. Journal of Educational Technology, 6(1), 1-7.

From 2009, Ege and Bayrakdaroğlu may be found. Analizi ve Lojistik Regresyon Tekniği ve Hisse Senedi Getiri Başarıların In MKB Şirketlerinin. Vol. 5, No. 10, pp. 139–158, ZKÜ Sosyal Bilimler Dergisi.

In 2005, Emel and Taşkın published a book. Ağaçları ve Ber Satış Analizi Uygulaması ve Veri Madenciliğinde Karar. In the Social Science Journal of Eskişehir Osmangazi University, volume 6, issue 2, pages 221-236. Written by Erdem in 2014. Geri Erişim ve Duygu Tanıma için Ses Sinyallerinde geri etkiler. Hizmet edilebilir, University of Başkent, Ankara.

The article was written by Gök (2017) and was published in the journal Academic Nursing. Journal of Gazi University Faculty of Science, 5(3), 139-148.

Gör, İ. (2014). Geometric Binary Search Algorithms for Data Retrieval and Analysis. Adnan Menderes University, Aydın, Turkey, is the birthplace of the author.

Mathematik 1 Derslerindeki Başarısın Destek Vektör Makineleri Kullanılarak Tahmin Edilmesi, Mühendislik Öğrencilerinin ćomak, E. (2011). The journal of Pamukkale University's Medical Sciences, volume 17, issue 2, pages 87–96.

By Güneren, H. in 2015. Yüz Tanıma Uygulaması Gömülü Sistem Üzerinde Destek Vektör Makineleri Kullanarak. I am a student at Yıldız Technical University in Istanbul.

Habiefendioğlu (2012). Teşhisi ve Makine Öğrenmesi Yöntemleri ile Glokom Hastalığının.

Faculty of Law, Selçuk University, Konya.

"Yapay Sinir Ağları ile Türkiye Elektrik Enerjisi Tüketiminin 2010 Yılına Kadar Tahmini" (The Energy Crisis and Its Impact on Turkey's Electric Power Sector, 2004). Published in the Gazi University Medical Faculty Journal, volume 19, issue 3, pages 227-233.

It was published in 2010 by Kavzoğlu and Çölkesen. The application of Destek Vektör Makineleri's Kernel Fonksiyonların Etkilerin İncelenmesi ile Uydu Görüntülerinin Sınıflandırılmasında. Vol. 144, Issues 73–81, Harita Dergisi.

This sentence is a citation for a 2016 publication by Kılınç, Borandağ, Yücalar, Tunalı, Şimşek, and Özçift. The K-Nearest Neighbors Algorithm and the R-Dili for Metin Data Quality Assessment. Vol. 28, No. 3, pages 89–94, Marmara Fen Bilimleri Dergisi.

Kızılkaya and Oğuzlar published a paper in 2018. Publication History of Denetimli Öğrenme Algorithms in R and Their Application. Karadeniz, 37(37), 90-98.

Kocadayı, Y., Erkaymaz, O., and Uzun, R. published this in 2017. The 81st Annual General Meeting of the Turkish Electricity Association. Research in Science and Technology: A Bilge International Journal, 1(1), 59-64.

The article "Öğrenci Notları ve Hastalık Verilerine Uygulanması Bitirme Tezi" was written by Özkan in 2013 and was published by İstanbul Teknik Üniversitesi in the field of K-Means Kümeleme and K-NN Support Vector Machines.

As said by Öztemel in 2003. Make a payment. Publication: Papatya Yayıncılık, Istanbul.

In 2016, Sağbaş and Ballı published a work. Ulaşım Türü Tespiti, Kadınların Telefonları ve Makine Öğrenmesi Kullanılarak. The journal of Pamukkale University's Medical Sciences, volume 22, issue 5, pages 376-383.

This information is from Sırmaçek (2007). The Fpga code is used for mobile robot control algorithms. I am a student at Yıldız Technical University in Istanbul.

This information is from Tantuğ and Türkmenoğlu (2015). Analyze the Turkish Meteors. The author is a student at İstanbul

Teknik Üniversitesi.

"Topal, Ç."—2017. Alan Turing's Top Secret Project: A Top Secret Project for You to Watch. the DTCF journal, volume 57, issue 2, pages 1350–1352.