NIPES

Journal of Science and Technology Research

Journal homepage: www.nipesjournals.org.ng

Solving Real-World Problems Using Machine Learning with Big Data

Omankwu, Obinnaya Chinecherem[1], Osodeke, Efe Charlse[2] and Kanu Chigbundu[3],
[1]Department of Computer Science. Michael Okpara University of Agriculture, Umudike, Nigeria.

Article Info

***Keywords*:** *machine learning, deep learning, big data, artificial intelligence*

Received 04 June 2023
Revised 02 July 2023
Accepted 09 July 2023
Available online 15 Sept. 2023
https://doi.org/10.5281/zenodo.8348943

ISSN-2682-5821/© 2023 NIPES Pub.

Abstract

Machine learning algorithms use big data to identify future trends and make predictions for your business. In an industry where understanding consumer patterns can lead to big improvements, machine learning is very efficient at deciphering data. Deploying machine learning can be a giant leap forward for an enterprise, and it's more than just integrating it at the top tier. This requires a redefinition of workflows, architecture, data collection and storage, analytics and other modules. The scope of the system overhaul should be assessed and clearly communicated to the appropriate parties. A major focus of machine learning is the development of computer programs that can access data and use it to learn. The learning process begins with observations or data to find patterns in the data and make better decisions. The main goal of data analysis using machine learning is for computers to automatically learn and adjust their behavior accordingly, without the need for human intervention or interaction. Considering the many applications where data analysis is found in the real world, this article will therefore focus on big data analysis and explore the fundamental applications of machine learning as one of the tools of artificial intelligence. The purpose of this article is to understand the aspects, components, applications, and challenges of introducing machine learning into the real world.

1. Introduction

Machine learning is the science of making computers learn and behave like humans. These systems automatically get smarter as they interact with data, networks, and people, ultimately becoming able to solve and predict the world's practical problems. Machine learning is a form of artificial intelligence that uses algorithms to extract patterns from raw data [1-10]. The main focus of machine learning is to enable computer systems to learn from experience without human intervention or explicit programming. There are many types of machine learning algorithms, and hundreds of new ones are created every day. They are usually categorized by learning style (supervised learning, unsupervised learning, semi-supervised learning, etc.) or their similarity [11]. They are formal and functional (e.g. classification, regression, decision trees, grouping, deep learning, etc.). Regardless of learning style or performance, all machine learning algorithms are: Representation: A set of classifiers or a language that the computer understands, Evaluation: Also known as objective performance/scoring, Optimization: search method; Often the classifier with the highest score. A fundamental goal of machine learning algorithms is to generalize insights beyond the trained examples i.e to successfully interpret the data.

In machine learning, rather than programming everything, data is fed to a general-purpose algorithm, which builds its own logic based on the data fed. There are various methods of machine learning such as supervised learning, unsupervised learning, and reinforcement learning. These three categories include algorithms used in machine learning. Machine learning has thus emerged as a new skill for computers. Today, this science is used in a variety of technical fields, and its use is so widespread that people are often unaware of its presence in everyday tools and gadgets.

Machine learning helps businesses understand their customers on a deeper level. Machine learning algorithms help teams align product development and marketing efforts with customer demand by collecting customer data and linking it to behavior over time. It's important to note that machine learning algorithms can learn from experience just like humans. As new data is input, these algorithms learn, change, and grow without changing code. The results obtained may not be very accurate initially, but machine learning algorithms can use their output data to improve future results [16,17].

It is difficult to imagine industrial activity without the use of machine learning and artificial intelligence. Machine learning is important due to its wide range of applications and tremendous adaptability, providing fast, effective and optimal solutions to complex problems [20,21]. Machine learning allows us to quickly and automatically generate models that can analyze larger and more complex data and deliver faster and more accurate results, even at very large scale. By creating accurate models, you are more likely to find profitable opportunities and avoid unknown risks [1, 2, 3,4]. Machine learning algorithms use a variety of techniques to manage large amounts of complex data for decision making. These algorithms take on the task of learning from data by providing specific inputs to the machine. It is very important to understand how these algorithms and machine learning systems work in order to know how to use them in the future. For these reasons, this book explores and analyzes different types of machine learning techniques and their real-world applications.

1.1. Types of machine learning

Machine learning is often classified according to how the algorithm learns. The four main approaches are:
Supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning. Each of these methods has a specific purpose. The type of algorithm a data scientist chooses depends on the type of data they want to predict. Of these, two main categories can be considered: supervised and unsupervised learning [13, 14, 15]. These two main types are shown in Figure 1.

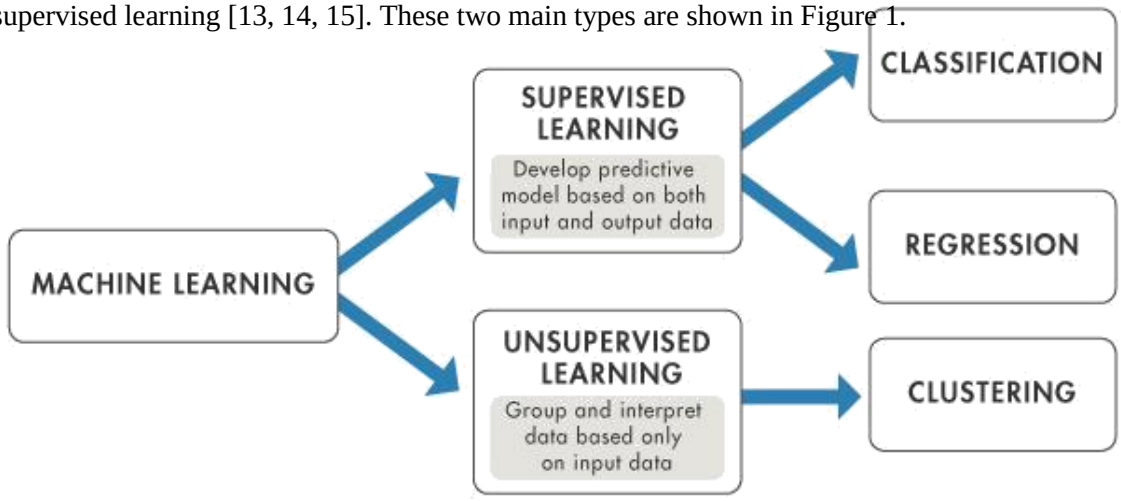

Fig. 1. Two main types of machine learning

1.2.1 Supervised learning

In supervised learning, a data scientist trains an algorithm using labeled data and variables that the algorithm is expected to evaluate. Once a model is trained on a set of known (labeled) data, unknown (unlabeled) data is fed into the model to get new answers. Supervised machine learning requires data scientists to train algorithms using both labeled inputs and desired outputs [13, 14, 15]. Figure 2 shows an overview of supervised machine learning.

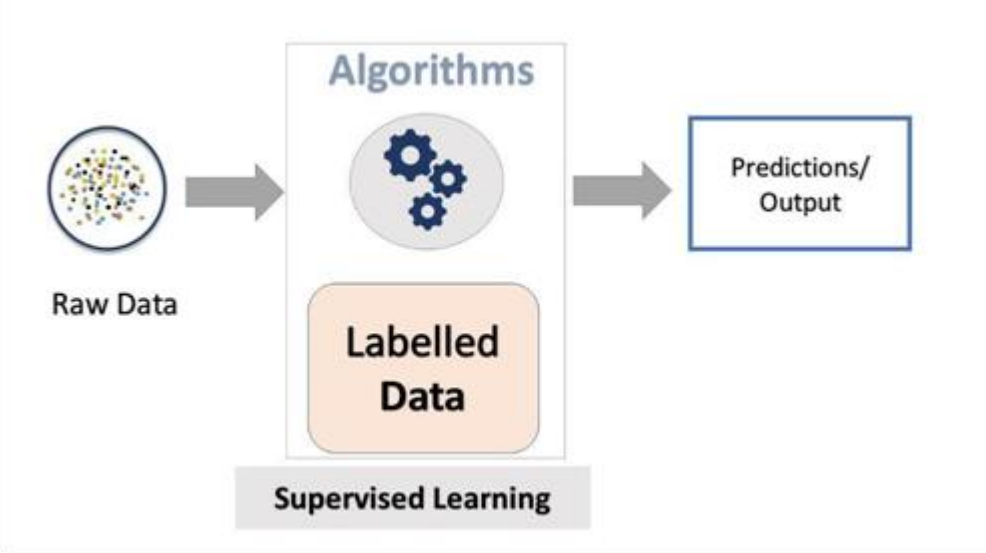

Fig. 2. Supervised machine learning [11, 12]

1.2.2. Unsupervised learning

This type of machine learning is an algorithm that learns from unlabeled data. Unsupervised machine learning algorithms examine unlabeled data looking for patterns that can be used to group data points into subsets. Note that unsupervised learning cannot add labels to the data [8].

Unsupervised learning algorithms are suitable for the following tasks:

1. Clustering: Dividing data sets into groups based on similarity.
2. Anomaly detection: identifying unusual data points in a data set.
3. Dependency exploration: Identifying sets of items in a data set that often occur together.
4. Dimensionality reduction: reducing the number of variables in a data set.

1.2.3. Semi-Supervised Learning

This approach combines supervised learning (training data is labeled) and unsupervised learning (training data is unlabeled). In semi-supervised learning, very little input data is labeled [19]. In semi-supervised machine learning, the model is first trained on labeled data. Unlabeled data is then entered into the model. The model labels unlabeled data with some degree of accuracy. This data is called pseudo-tagged data. Finally, we use a combination of the pseudo-tagged data and the original tagged data to improve the accuracy of the model.

Today, there is an enormous amount of data in various industries. Many large companies have collected, and are still collecting, millions of terabytes of data. However, indexing the collected data is very expensive as it requires a lot of staff and resources. A large number of real databases fall into this category [8].

Some of the fields in which semi-supervised learning is used are:

[1] Machine translation: teaching language translation algorithms.

[2] Fraud detection: Identifying fraud cases when there are only a few positive samples.

[3] Data Labeling: Algorithms trained on small data sets can learn to automatically apply data labels to larger data sets.

1.2.4. *Reinforcement learning*

Data scientists typically use reinforced machine learning to perform multi-step processes with specific rules. A data scientist programs an algorithm to perform a task and gives a positive or negative signal when the algorithm attempts to do it. If the model predicts or produces an outcome, it is penalized if the prediction is wrong and rewarded if the prediction is correct, training the model accordingly [11, 12].

Some of the fields in which reinforcement learning is used are:

1. Robotics: Robots can learn to perform real-world tasks using this technique.
2. Video games: Reinforcement learning has been used to train robots to play a number of video games.
3. Resource management: Given limited resources and a specific goal, reinforcement learning can help companies plan how to allocate resources.

1.3 Machine learning processes

Explaining how complex machine learning models work can be difficult. Some industries require data scientists to use simple models [7]. Because it's important to explain how this company's decisions are made. Sophisticated models can provide accurate predictions, but it is difficult for laymen to explain how the output is determined. In general, the machine learning process is:

- Data collection and processing

The first step in the machine learning process is to provide the machine with the necessary knowledge and data. This data is split into two groups. One group is used for training the system and the other group is used for testing the system [6].

- Model selection and training

The second next step in machine learning principles is choosing and training a model. There are many different types of machine learning algorithms and models that have already been developed and modified to solve specific types of problems. Therefore, we select and train models according to their need and suitability for solving a particular problem [9].

- Model evaluation

Machines learn different patterns and characteristics from the data they are taught and train themselves to make decisions in different areas such as identifying, classifying and predicting new data. To check exactly how well the machine makes these decisions, we test our predictions against training data [23].

1.4. Machine learning applications

Machine learning is everywhere. Even though you use it many times in your daily life, you may not know it. Key uses of machine learning are [18].

- Image recognition and processing

One of the most popular uses of machine learning is image recognition. There are many situations to classify objects in digital images, and machine learning can be used for this purpose.

- Voice recognition

Another application of machine learning is converting speech to text. This problem is also known as Computer Speech Recognition with Automatic Speech Recognition.

- Diagnosis

Machine learning can be used in techniques and tools for diagnosing disease. With this technology, clinical parameters can be analyzed and combined to predict disease progression, extract medical information, conduct outcome-driven research, and implement treatment

- Statistical analysis

One of the most important problems in economic problems is finding a short-term strategy for buying and selling stocks. To obtain these strategies, users use trading algorithms to buy and sell securities based on factors such as historical correlations and general economic variables. Machine learning is very helpful in obtaining these short-term strategy algorithms.

- Learning community

A learning community is the process of developing knowledge across different products. One use of machine learning is to study the relationships between the products people buy.

- Classification

The importance of classification is to classify each person, object, etc. into different categories as a subject of study. Classification helps analyze an object's measurements to identify the category to which it belongs. Analysts use data to build effective relationships [24].

Machine learning has many capabilities that can be used in finance and banking. Popular financial services can be provided using machine learning and artificial intelligence. Machine learning can help banks and financial institutions make more accurate decisions, help financial service providers proactively spot account closures, track customer spending behavior, conduct market analysis, and identify patterns. , teach intelligent machines the cost, and finally allow machine learning algorithms to: You can easily recognize upcoming trends and react in real time [11, 12, 13, 14].

Machine learning algorithms have been around for decades, but have gained new popularity with the advent of artificial intelligence. Deep learning models in particular power today's cutting-edge AI applications. The battle for machine learning platforms is intensifying as machine learning becomes more important to business operations and artificial intelligence becomes more applicable in enterprise environments [11, 12, 13, 14].

The field of machine learning is a tool for improving and adapting to the pressing problems of today's world, making our lives easier and more modern. The main difference between old software and machine learning-based software is that in the new method the system is trained to use large amounts of information and this method is also based on experience. This makes machine learning approaches significantly more effective than older approaches to problem solving. For this reason, most the study evaluated important applications of machine learning.

3. Conclusion

As mentioned earlier, machine learning is a subset of artificial intelligence. Through the use of machine learning techniques, computer systems can learn patterns in the information or data being processed and make use of them. Also note that these techniques teach computer systems without explicit programming.

Machine learning solutions continue to revolutionize core business processes and are increasingly used in our daily lives. Many companies have already started using machine learning as it has the potential to enable more accurate predictions and business decisions. Continued research in deep learning and artificial intelligence is increasingly focused on developing more general applications. Today's AI models require extensive training to create highly optimized machine learning algorithms to perform their tasks. However, some researchers are exploring ways to make the model more flexible, looking for techniques that would allow the machine to apply the context it learned from her one task to various future tasks.

References

[1] Aliahmadi, A., & Nozari, H. (2023). Evaluation of security metrics in AIoT and blockchain-based supply chain by Neutrosophic decision-making method. In Supply Chain Forum: An International Journal (Vol. 24, No. 1, pp. 31-42). Taylor & Francis.
[2] Aliahmadi, A., Nozari, H., & Ghahremani-Nahr, J. (2022). A framework for IoT and Blockchain Based on Marketing Systems with an Emphasis on Big Data Analysis. International journal of Innovation in Marketing Elements, 2(1), 25-34.
[3] Aliahmadi, A., Nozari, H., & Ghahremani-Nahr, J. (2022). AIoT-based sustainable smart supply chain framework. International journal of innovation in management, economics and social sciences, 2(2), 28-38.
[4] Aliahmadi, A., Nozari, H., & Ghahremani-Nahr, J. (2022). Big Data IoT-based agile-lean logistic in pharmaceutical industries. International Journal of Innovation in Management, Economics and Social Sciences, 2(3), 70-81.
[5] Bayanati, M., Peivandizadeh, A., Heidari, M. R., Foroutan Mofrad, S., Sasouli, M. R., & Pourghader Chobar, A. (2022). Prioritize Strategies to Address the Sustainable Supply Chain Innovation Using Multicriteria Decision-Making Methods. Complexity, 2022.
[6] Chen, Y., Bayanati, M., Ebrahimi, M., & Khalijian, S. (2022). A Novel Optimization Approach for Educational Class Scheduling with considering the Students and Teachers' Preferences. Discrete Dynamics in Nature and Society, 2022.
[7] Eisapour, K., Bayanati, M., & Yousefpour, J. (2013). A mathematical model for ranking R&D organizationsas a technology development factor. Advances in Environmental Biology, 717-721.
[8] Ghahremani-Nahr, J., & Nozari, H. (2021). A Survey for Investigating Key Performance Indicators in Digital Marketing. International journal of Innovation in Marketing Elements, 1(1), 1-6.
[9] Moeini, E., Bayanati, M., Givi, S. B., & Soheili, S. (2013). Individual social responsibility a platform for intellectual property rights: Iranian users of CDs case. Advances in Environmental Biology, 709-717.
[10] Nahr, J. G., Nozari, H., & Sadeghi, M. E. (2021). Green supply chain based on artificial intelligence of things (AIoT). International Journal of Innovation in Management, Economics and Social Sciences, 1(2), 56-63.
[11] Nozari, H., & Nahr, J. G. (2022). The Impact of Blockchain Technology and The Internet of Things on the Agile and Sustainable Supply Chain. International Journal of Innovation in Engineering, 2(2), 33-41.
[12] Nozari, H., & Sadeghi, M. E. (2021). Artificial intelligence and Machine Learning for Real-world problems (A survey). International Journal of Innovation in Engineering, 1(3), 38-47.
[13] Nozari, H., & Szmelter-Jarosz, A. (2022). IoT-based Supply Chain For Smart Business (Vol. 1). ISNET.
[14] Nozari, H., Fallah, M., Kazemipoor, H., & Najafi, S. E. (2021). Big data analysis of IoT-based supply chain management considering FMCG industries. Бизнес-информатика, 15(1 (eng)), 78-96.
[15] Nozari, H., Fallah, M., Szmelter-Jarosz, A., & Krzemiński, M. (2021). Analysis of security criteria for IoT-based supply chain: a case study of FMCG industries. Central European Management Journal, 29(4).
[16] Nozari, H., Ghahremani-Nahr, J., & Szmelter-Jarosz, A. (2023). AI and machine learning for real-world problems. Advances In Computers, (online first).
[17] Nozari, H., Szmelter-Jarosz, A., & Ghahremani-Nahr, J. (2022). Analysis of the Challenges of Artificial Intelligence of Things (AIoT) for the Smart Supply Chain (Case Study: FMCG Industries). Sensors, 22(8), 2931.
[18] Pishkar, N., Nasimi, M. A., & RAHMATI, M. (2021). Developing a Conceptual Model of Green Supply Chain Antecedents and Consequences in the Qualification Approach.
[19] Rafierad, S., Aghajani, H. A., Agha Ahmadi, G., & Rahmaty, M. (2022). Construction and Validation of Dimensions and Components of the Organizational Anomie Scale in order to provide a Native Model in Government Hospitals. Journal of System Management, 8(2), 57-73.
[20] Shayan Nia, S. A., & Mirataollahi Olya, R. (2021). Provide a model for manpower scheduling using a hybrid meta-innovative algorithm in the Water and Sewerage Company. Journal of Industrial Strategic Management, 6(2), 18-32.
[21] Shayannia, S. A. (2022). Designing a Multiobjective Human Resource Scheduling Model Using the Tabu Search Algorithm. Discrete Dynamics in Nature and Society, 2022.
[22] Tavakkoli-Moghaddam, R., Ghahremani-Nahr, J., Samadi Parviznejad, P., Nozari, H., & Najafi, E. (2022). Application of internet of things in the food supply chain: a literature review. Journal of applied research on industrial engineering, 9(4), 475-492.
[23] Toloie-Eshlaghy, A., & Bayanati, M. (2013). Ranking information system success factors in mobile banking systems with VIKOR. Middle-East Journal of Scientific Research, 13(11), 1515-25.
[24] Tootian, S., Bayanaty, M., & Jalali, A. (2022). Identify The Dimensions And Components Of Communication Technology In Order To Empower Women Entrepreneurs In Business. An Approach to Business Management, 2(4), 32-49.