UNIVERSITY OF WESTMINSTER

INFORMATICS INSTITUTE OF TECHNOLOGY

Informatics Institute of Technology

Trends in Computer Science

4COSC008C

Machine Learning

Overwiew of Machine learning. Describe and compare two different Machine Learning Techniques

Baminahannadige Oshadha Kumarinda Fernando

W2052726

20220857

Group Members

1. Sihath Senarath Yapa w2052733
2. Prashan Andradi w2052718
3. Raveen Kalsara w2053781
4. Imeth Rajapaksha w2052725

Contents

Table of Figures

1. OVERWIEW OF MACHINE LEARNING.

Machine learning is a subfield of Artificial Intelligence that uses algorithms to learn from vast datasets. It is used in various applications, such as virtual personal assistants like Siri, Alexa, and Google Assistant, social media platforms like Instagram and Facebook, professional networking sites like LinkedIn, chatbots powered by Natural Language Processing (NLP), email spam filtering, voice recognition technology, and fraud detection. Machine learning algorithms are essential for enhancing user experience, security, and enhancing communication with devices. As we move into the digital age, the influence and applications of machine learning are expected to expand, offering endless possibilities.

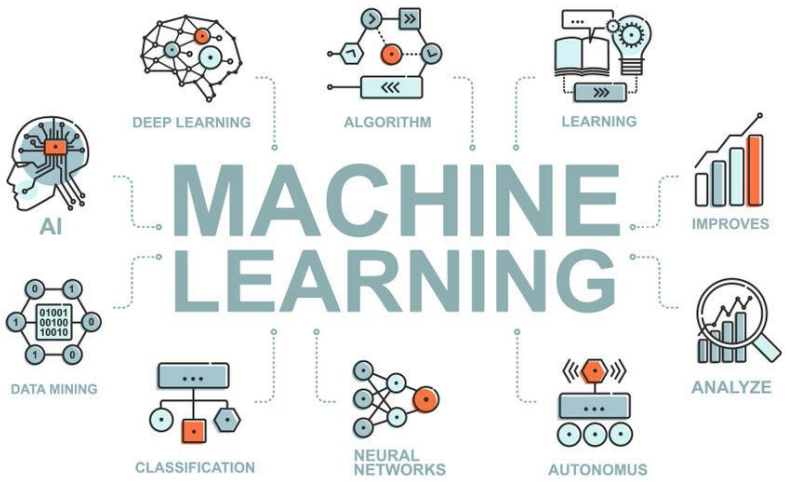

Figure 1Machine Learning

2. MACHINE LEARNING TECHNIQUES.

The world is getting “smarter” every day, and business are utilizing machine learning algorithms more frequently to streamline process in order to meet customer expectations. They can be found in end-user devices or in credit card fraud detection. Supervised learning and unsupervised learning are two basic techniques in artificial intelligence (AI) and machine learning. The primary distinction is that one makes use of labeled data to aid in result prediction, while the other does not.

2.1 Supervised learning

A subset of artificial intelligence and machine learning is called supervised learning, or supervised machine learning. Its use of labeled datasets to train algorithms for precise data classification or result prediction defines it. The cross-validation procedure involves the model adjusting its weights as input data is fed into it until the model has been fitted adequately. Classifying spam in a different folder from your inbox is one example of the many real-world problems that supervised learning helps enterprises solve at scale. Similar attention is made in dealing NTL detection through supervised learning techniques. For example, Zheng et al. [21] have experimented wide and deep con volutional neural network (CNN) in a dataset collected from a Chinese electricity company. Khawaja Moyeezullah Ghori, Muhummadimran (December 26, 2019)

A training set is used in supervised learning to teach models to produce the results. The model can learn over time thanks to the inputs and accurate outputs int this training dataset. Through the loss function, the algorithm gauges its accuracy and adjust until the error is suitably reduced. When using data mining, supervised learning may be divided into two categories of problems they are regression and classification.

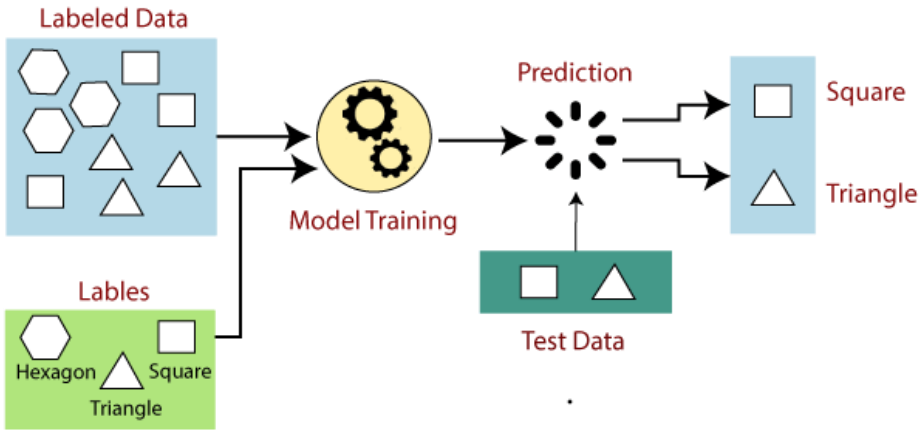

Figure 2 : Supervised Learning

An algorithm is used in classification to precisely place test data into designated groups. It identifies particular entities in the dataset and tries to make recommendations on the definition or labelling of those items. The next section provides a more detailed description of several popular classification techniques, including random forest, k-nearest neighbor, decision trees, support vector machines (SVM), and linear classifiers. The cross-validation procedure involves the model adjusting its weights as input data is fed into it until the model has been fitted adequately. Classifying spam in a different folder from your inbox is one example of the many real-world problems that supervised learning helps enterprises solve at scale.

2.1.1 Advantages

Linear regression is a simple and adaptable method that can be easily updated with new data using stochastic gradient descent. Supervised learning offers more precise results and can solve various linear and non-linear problems efficiently, including robotics, prediction, classification, and industrial control.

2.1.2 Disadvantages

Supervised learning is time-consuming and underperforms in complex situations due to its complexity and adaptability. It is not cost-effective for large datasets or when data labelling increases, as manual labelling is expensive and the best image dataset has only one thousand labels.

2.2 Unsupervised learning

Unsupervised learning in artificial intelligence refers to machine learning that occurs without human supervision. Unsupervised machine learning models, in contrast to supervised learning, are given unlabelled data and let to find patterns and insights on their own—without explicit direction or instruction.

Unsupervised learning, as the name implies, makes use of self-learning algorithms that acquire knowledge without the need for labels or prior training. Rather of receiving explicit instructions on how to operate with each piece of data, the model is given raw, unlabeled input and is forced to deduce its own rules then arrange the data according to patterns, distinctions, and similarities. Algorithms for unsupervised learning are more appropriate for more intricate processing jobs, including clustering big datasets.

2.2.1 Advantages

Unsupervised learning is widely used in real-time to uncover hidden patterns in industries, potentially leading to new business ventures. It offers less complexity due to the absence of label interpretation and easier data acquisition.

2.2.2 Disadvantage

Unsupervised learning has drawbacks such as increased cost due to human intervention, uncertainty in utility due to lack of label or output measure, unclear output and sorting, and lower precision in results.

3. COMPARING BETWEEN SUPERVISED MACHINE AND UNSUPERVISED MACHINE LEARNING.

The two methods of machine learning are supervised and unsupervised learning. However, the two methods are applied to distinct situations and datasets. Learning under supervision is similar to having a tutor walk you through a subject with precise directions and examples. Given labeled data, you are predisposed to knowing the right answers. The objective is to train a model so that it can correctly forecast results when presented with new data.

It's similar to studying from a textbook when all the solutions are given. Conversely, unsupervised learning is more akin to discovering a new city unaccompanied by a map or tour guide. Unlabeled data is provided to you, so there are no predetermined correct or incorrect responses. Rather, the objective is to find hidden structures or patterns in the data. It's similar to attempting to find fascinating locations in a city without assistance.

Like a student getting feedback from a teacher, you regular compare your progress in supervised learning against the right response. Unsupervised learning however, lacks this kind of feedback loop

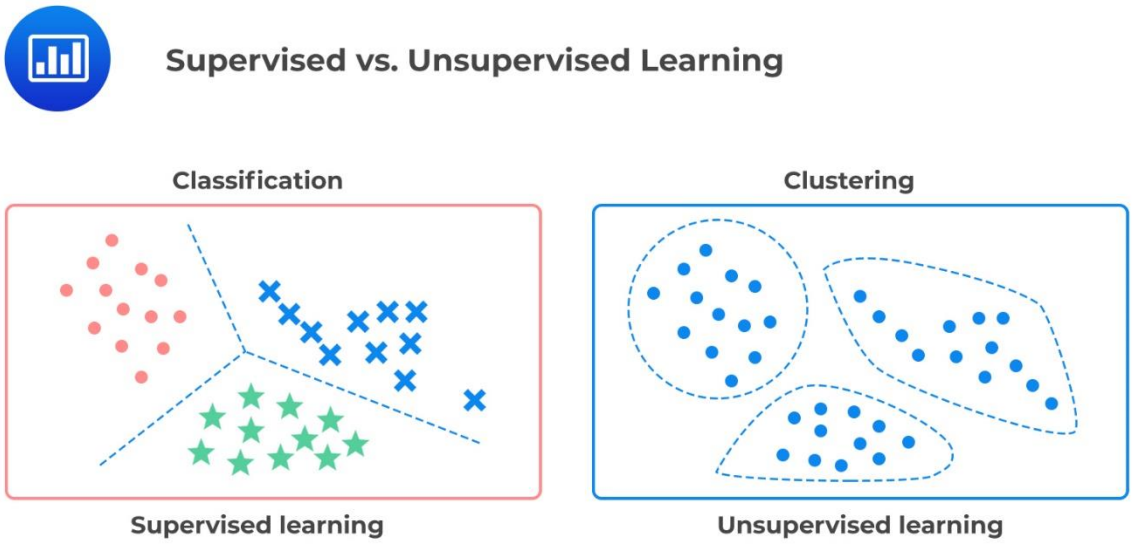

Figure 3Supervised vs Unsupervised

When you have distinct input-output pairings, such as when estimating home values based on attributes like size and location, supervised learning works really well. Unsupervised learning, on the other hand, is helpful when all you have is input data and you want to sift through it to find insights, such as distinguishing various client segments according to their purchasing

patterns.

Unsupervised learning is more similar to actual artificial intelligence, while supervised learning typically produces more accurate results because it is guided by known consequences.

4. CONCLUSION

Finally, a comparison of supervised and unsupervised learning strategies shows the different functions and uses of each in the field of machine learning. Supervised learning uses labeled data to train models for accurate prediction tasks, much like a guided lesson. On the other hand, unsupervised learning explores unexplored areas by independently identifying latent structures and patterns in unlabeled datasets. Unsupervised learning is superior at extracting insights from unprocessed data without predetermined results, whereas supervised learning performs best in jobs with explicit input-output correlations.

Both methods, which are appropriate for various situations and goals, provide significant advances in AI and machine learning capabilities. The decision between supervised and unsupervised learning ultimately comes down to the type of data and the intended result, emphasizing how critical it is to recognize and capitalize on each met each advantage for the best outcomes

5. REFERENCES

Andreas C. Muller, S.G. (2016).*Introduction to Machine Learning with Python [Book]*. [online] www.oreilly.com. Available at: https://learning.oreilly.com/library/view/introduction-to-machine/9781449369880/ch02.html [Accessed 20 Feb. 2024].

Khawaja Moyeezullah Ghori, Muhummadimran .(2019). Performance Analysis of Different Types of Machine Learning Classifiers for Non-Technical Loss Detection .Available from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8943419 [Accessed on February 25,2024]

Mehrabi, N. et al. (2021) "A Survey on Bias and Fairness in Machine Learning", ACM Computing Surveys, 54(6), pp. 1–35. Available at: https://dl.acm.org/doi/10.1145/3457607 [Accessed on February 24, 2024].

Raschka, S. (2019). Python Machine Learning. Available from https://books.google.lk/books?id=n1cjyAEACAAJ&dq=%22Python+Machine+Learning%22+by+Sebastian+Raschka+and+Vahid+Mirjalili&hl=en&sa=X&redir_esc=y [Accessed 24 February 2024].

Sams, J. (no date). Risks of Machine Learning. Javatpoint. Available from Risks of Machine Learning - Javatpoint [Accessed 25 February 2024].

Sudipta, M. (2021). ML.NET : Machine Learning to your Applications. 1st ed. Available from ML.NET revealed : simple tools for applying machine learning to your applications - University of Westminster [Accessed 22 February 2024].