Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 216 (2023) 565–570

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Computer Science and Computational Intelligence 2022

General pattern recognition using machine learning in the cloud

Albert Salim[a], Juliandry[a], Louis Raymond[a], Jurike V Moniaga[a,*]

[a]*Computer Science Depertment, Bina Nusantara University, Jakarta 11530, Indonesia*

Abstract

Machine learning (ML) and cloud computing are two subjects that mix very well. The existence of cloud computing enables data scientists to create their machine learning models with the benefits of cloud computing which are very low cost, high performance, and high availability. This opens a new opportunity of allowing students and other people to utilize powerful machines for daily use with the help of appropriate cloud service. On the other hand, machine learning is usually perfect for recognizing patterns in our daily life. This article discusses the possibility and the benefit of creating a user-friendly general-purpose pattern recognition that can be played around by anyone using software as a service.

Keywords: cloud computing; machine learning; software as a service; machine learning as a service; pattern recognition

1. Introduction

1.1. Introduction

Pattern Recognition is one of the most well-known uses of Machine Learning. We find many patterns in our everyday lives, like differentiating by colors and shapes. With the help of Machine Learning, we are able to recognize patterns that even our brains find hard to think about. However, machine learning requires computing power [1]. Acquiring a dedicated high-performance computer and maintaining it would be costly. Therefore, cloud computing idea was realized and it is now used globally, ranging from personal to enterprise applications.

* Corresponding author.
E-mail address: jurike@binus.edu

1877-0509
10.1016/j.procs.2022.12.170

Cloud computing enables people to use their internet connection for any computation without having to interact physically with the hardware. Its services can be considered pricey for some use cases or way cheaper [2] than having to own and manage your own server with computational power of similar grade, hence being a good consideration for anyone, especially server admins and data scientists who may find high availability, extra computational power that can be shut down anytime, and the feature of not needing to manage the servers by themselves [3]. As of today, cloud computing in the Industrial Revolution 4.0 is utilized for several purposes, such as Big Data Analysis [4] and Internet of Things [5][6].

Nevertheless, the implementation of an intuitive and robust pattern recognition goes a long way, from fine-tuned parameters to sufficient amounts of well-conditioned datasets. Development of a machine learning model may take years of dedicated research just to reach a few percent better accuracy, considering challenges such as patterns that may vary from one dataset to another [7] . Also, the concept of pattern recognition itself is still thought to be complex and not many realize that some of the devices they use regularly already implements machine learning[8].

With the ongoing trend in studies using Machine Learning and Cloud computing in the form of Machine Learning as a Service (MLaaS) for several purposes, this paper will showcase what we have known so far and analyze whether it is feasible to develop a general-purpose pattern recognition application that is accessible and usable to common people. The main difference between this paper and other similar work referenced is the proposed idea of a pattern recognition application which aims towards a more flexible implementation compared to most research papers with specific purposes.

1.2. Objectives

For this research paper there are objectives that are made to define the scoping of the topic and as a guideline for this paper in general: propose a general pattern recognition using Cloud Computing, evaluate similar research that has been done before and review said papers for future learning, and achieve a better understanding about pattern recognition using Machine Learning combined with Cloud technologies.

1.3. Methodology

To satisfy the given objectives, we will perform a literature review to evaluate previous research from the past 5 years with relevant topics. Additional literature about Machine Learning and Cloud Computing will also be used to add prior knowledge of the topic.

2. Literature Review

2.1. Machine Learning (ML) [9]

a Machine Learning is a technique in computer programming to enable machines to "learn" from experience or example. This technique applies a model to set rules which work best for mapping problems to the supposed solution. For example, in Computer Vision (CV) for facial recognition, one way of building its machine learning model is by using linear regression, which applies the following equation:

$$f(x) = W \times x \tag{1}$$

$\mathrm{W} = \mathrm{parameter(weight\ vector)}$
$\mathrm{x} = \mathrm{input}$

2.2. Pattern Recognition

Pattern recognition is the study of understanding data in order to extract meanings and features inside patterns. Patterns can be found everywhere if we carefully look to our surroundings. They are generally used by us to identify any kind of thing. Portraying us, machines can also take advantage of patterns for any possible purpose [10].

2.3. Cloud Computing

Cloud computing is a concept where computer services live and are available over the internet. This makes server owners easily able to run their server without any expertise of computer server hardware and software management and maintenance. Cloud computing also gives server administrators the power to easily rent a scalable and pay-what-you-use server [11]. A Cloud is a server pool which can provide computing resources to their clients. There are a few characteristics of Cloud Computing such as Self-Healing, Multi-tenancy, Linearly Scalable, Service-oriented, SLA Driven, and Flexible

There are four essential elements in Cloud Computing: on-demand self-service, where users can request for computing resources without having to contact the providers directly for permission; broad network access, where the computing resources are delivered across a network to multiplatform users; resource pooling [12], describes how the resources are 'pooled' together in a single location and cause physical computing resources to seem 'invisible' to users; and measured service, used to measure resource usage for payment. The cloud community also uses the following service models to categorize the Cloud Service : Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Data storage as a Service (DaaS)[13].

Cloud Service categories reflect to what service is offered by the cloud computing provider. Infrastructure as a service offers the whole machine to be used by the client however they want. IaaS usually uses virtualization technology named virtual machine, but a physical machine would theoretically be included as infrastructure as a service. IaaS is virtual machine rental over the internet. Platform as a service offers the building block for deploying cloud applications. PaaS usually uses a containerized cloud environment for its infrastructure, however PaaS itself is a managed environment for cloud applications to be deployed. PaaS examples are easy setup game server and one-click blog server builder available on the internet. Software as a service offers an application ready for end users to use. SaaS runs on top of any server that is visible to a network. SaaS applications usually are applications that are similar to desktop applications. SaaS allows the clients to enjoy their applications over the web, potentially integrating all of their devices [14]. Data storage as a Service, usually called cloud storage, acts as the storage on the internet which can be used for servers or personal use by mounting the cloud storage via SFTP, FTP, SMB, WebDAV, or other methods. [15] The world of cloud computing as a service does not stop there. There are various other kinds of cloud computing as a service that is not as popular if called with "as a service" term or rather circled terms. For example, there are gaming as a service which is usually called cloud gaming which allows gaming from the cloud, printing as a service which allows printing from distant locations, and backend as a service which allows app developers to manage their backend application from the cloud.

Modern cloud computing is known to be very close to "as a service" model, that there is an idea of making everything as a service (EaaS). Anyone can have a very powerful machine just for an hour which doesn't break the bank if compared to buying the machine instead with its maintenance and electricity bill. If the user wants to use the server just like a full traditional computer, IaaS is there for them. If the user wants to use the server for a more focused purpose, there is also PaaS. If the user just wants their life easier with a more available anywhere and anytime kind of app on the web, they can use SaaS. And if the user needs to have the flexibility of IaaS but the feature of PaaS, container-based system also comes as a service. This kind of system takes advantage of the namespace feature that Linux kernel has. Namespace allows server applications to run without conflict and take lower security and privacy risk.

2.4. Machine Learning as a Service (MLaaS)

An MLaaS platform is a ML platform that uses cloud to provide an ML service remotely where users can train and implement ML models. In some aspects. MLaaS is considered cheaper and more flexible than its offline counterpart due to its cloud-based infrastructure. MLaaS is accessed by a web page interface, and some of them use SaaS for the frontend to help even inexperienced users to access the MLaaS [9].

Most big companies that provide Cloud Services also have MLaaS [16]. A few examples would be Google, Amazon[17], and Microsoft [18][19]. These MLaaS platforms help their users to train ML models online using APIs to receive data remotely and to train the models. They also give ease to users to either distribute the trained model by using a certain API or to transfer it to the user's repository. Fig. 1. describes the typical steps taken by the user

while using user-managed Machine Learning service. We can also see that there are some steps that are not controllable by the user.

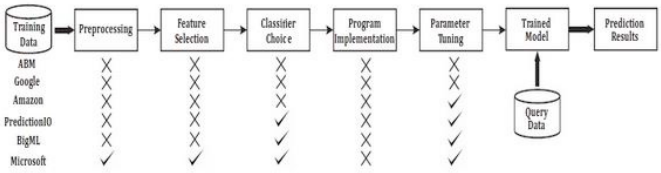

Fig. 1. MLaaS Pipeline and its user-controllable aspects.

3. Pattern Recognition with Maching Learning

This section explains how each machine learning method can be used for pattern recognition while considering its advantages and limitations to determine which method suits the best for pattern recognition.

3.1. Supervised Learning

Supervised learning is an approach in Machine Learning where an ML model is given a set of data with its own unique characteristics called features, and a "name" or an identifier of which class each data belongs to called labels. This dataset with labels is then used to train the Machine Learning model so that it would be able to predict the label of a new foreign test data that is related to the previous given dataset.

A few most known implementations of supervised learning are classification and regression [20]. In terms of pattern recognition, this approach gives a predictable outcome but may not be the optimal solution due to its risk of remembering the noise on the training data and causing overfitting [21].

3.2. Unsupervised Learning

In unsupervised learning, instead of training a model with data and making it guess between a certain number of available classes, the machine simply receives inputs and makes conclusions based on which algorithm is used. The conclusions are in the form of clusters filled with data points with similar values. These conclusions may then be used to group data inputs, to classify future inputs, or maybe as a part of preprocessing to be fit to another ML model.

Unsupervised learning utilizes the machine's ability to find patterns in data that may even be incomprehensible to humans. Some of the implementations of unsupervised learning are related to dimensionality reduction, where multiple-dimensional data are processed so that it could be presented in a much more intuitive two or three dimensional visualizations. In practice, this approach does not require too much time for gathering training data, but it may result in unpredictable outcomes that even researchers may not expect.

3.3. Reinforcement Learning

Reinforced learning holds the rule of training an algorithm (machine learning model) by giving it reward or punishment, depending on the result which is output by the machine to indicate what the machine should do and should not do. This approach mimics how we teach children by giving them treats if they behave properly. The difference is instead of giving the children treats, the machine is given a numerical value as its result score, which the machine is programmed to obtain however way it is, to maximize its total score from [22].

4. Works Related to Our Topic

During the process of Literature Review, we found a few proposals or papers that discuss topics that may be related to what we had in mind. We consider these researches associated to some extent with our proposed idea and could use these references to enhance our comprehension on the raise topic:

4.1. Biometric Pattern Recognition [22]

The paper written by Nicolas Ortiz, Ruben Dario Hernandez, Robinson Jimenez, Mauricio Mauledeoux, and Oscar Aviles describes how distinct patterns gathered from various parts of the human body can be recognized using Machine Learning algorithms. It uses Supervised, Unsupervised, and Reinforcement Learning methods to identify a given data from a dataset and perform classifications.

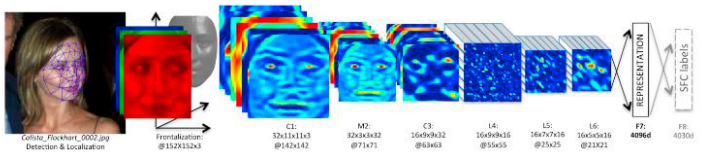

Fig. 2. Outline of a face depth architecture used in supervised learning

4.2. Using Pattern Recognition to RFID [23]

RFID or more commonly known as Radio Frequency Identification is a method of using a tag that is attached to an object that can be detected by a radio frequency for the purpose of tracking and identifying the object. Because this tag can be detected, then the pattern in the signal it sent can be extracted. The signal pattern / frequency that the RFID reader scans is sent to a cloud database, so that it can be normalized and fed into a pattern recognition model. This model is used for training from the data in the database and predicting new patterns that come in whether it is similar to which group of data. In this case the output of the model is a tagID(alphabet). The tagID / output is arranged by alphabets because it is known to resonate a certain frequency that can be read by the RFID reader. So the model could predict whether the signal pattern / frequency is close to which tagID. The better the accuracy from the training data, the better the model when predicting new data. The examples of pattern recognition used in the models are KNN (K-Nearest Neighbour) and SVM (Support Vector Machine)

References

[1] Thompson, N. C., Greenewald, K., Lee, K., & Manso, G. F. (2020). The computational limits of deep learning. arXiv preprint arXiv:2007.05558.

[2] Hajimirzaei, B., & Navimipour, N. (2019). Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. ICT Express, 5(1), 56-59. doi: 10.1016/j.icte.2018.01.014

[3] Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2016). Big Data and cloud computing: innovation opportunities and challenges. International Journal Of Digital Earth, 10(1), 13-53. doi: 10.1080/17538947.2016.1239771 --

[4] Wu, C., Buyya, R., & Ramamohanarao, K. (2016). Big data analytics= machine learning + cloud computing. arXiv preprint arXiv:1601.03115

[5] Santur, Y., Karaköse, E., Karaköse, M., & Akın, E. (2017). An Artificial Management Platform Based on Deep Learning Using Cloud Computing for Smart Cities. International Journal of Applied Mathematics, Electronics and Computers, 5, 24-28. doi: 10.18100/ijamec.2017SpecialIssue30466

[6] Bacciu, D., Chessa, S., Gallicchio, C., & Micheli, A. (2017, October). On the need of machine learning as a service for the internet of things. In Proceedings of the 1st International Conference on Internet of Things and Machine Learning (pp. 1-8). doi: 10.1145/3109761.3109783

[7] Biggio, B., & Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. Pattern Recognition, 84, 317-331. doi: 10.1016/j.patcog.2018.07.023

[8] Ipsos, M. O. R. I. (2017). Public views of machine learning. Report title, 92.

[9] Yao, Y., Xiao, Z., Wang, B., Viswanath, B., Zheng, H., & Zhao, B. Y. (2017, November). Complexity vs. performance: empirical analysis of machine learning as a service. In Proceedings of the 2017 Internet Measurement Conference (pp. 384-397). doi: 10.1145/3131365.3131372 --

[10] Learning, P., Bishop, C., & York, S. (2020). Pattern Recognition and Machine Learning | Christopher Bishop | Springer. Retrieved 14 April 2020, from https://www.springer.com/gp/book/9780387310732 --

[11] Wu, C., Nadjaran Toosi, A., Buyya, R., & Ramamohanarao, K. (2018). Hedonic Pricing of Cloud Computing Services. IEEE Transactions On Cloud Computing, 1-1. doi: 10.1109/tcc.2018.2858266 --

[12] Moreno-Vozmediano, R., Montero, R., Huedo, E., & Llorente, I. (2019). Efficient resource provisioning for elastic Cloud services based on machine learning techniques. Journal Of Cloud Computing, 8(1). doi: 10.1186/s13677-019-0128-9 --

[13] Yang, C., Yu, M., Hu, F., Jiang, Y., & Li, Y. (2017). Utilizing Cloud Computing to address big geospatial data challenges. Computers, Environment And Urban Systems, 61, 120-128. doi: 10.1016/j.compenvurbsys.2016.10.010 --

[14] van de Weerd, I., Mangula, I., & Brinkkemper, S. (2016). Adoption of software as a service in Indonesia: Examining the influence of organizational factors. Information & Management, 53(7), 915-928. doi: 10.1016/j.im.2016.05.008

[15] Juee U.Daryapurkar, P. (2014). Cloud Computing Issues and Challenges. International Journal On Recent And Innovation Trends In Computing And Communication, 2(4), 770-773. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/3057 --

[16] Hesamifard, E., Takabi, H., Ghasemi, M., & Jones, C. (2017, November). Privacy-preserving machine learning in cloud. In Proceedings of the 2017 on Cloud Computing Security Workshop (pp. 39-43). doi: 10.1145/3140649.3140655 --

[17] Kim, H., Kim, M., Seo, D., Kim, J., Park, H., Park, S., ... & Sung, N. (2018). Nsml: Meet the mlaas platform with a real-world case study. arXiv preprint arXiv:1810.09957.--

[18] Hunt, T., Song, C., Shokri, R., Shmatikov, V., & Witchel, E. (2018). Chiron: Privacy-preserving machine learning as a service. arXiv preprint arXiv:1803.05961.--

[19] Subbiah, U., Ramachandran, M., & Mahmood, Z. (2019, January). Software engineering approach to bug prediction models using machine learning as a service (MLaaS). In ICSOFT 2018-Proceedings of the 13th International Conference on Software Technologies (pp. 879-887). doi: 10.5220/0006926308790887 --

[20] Bishop, C. M. (2006). Pattern recognition and machine learning. springe --

[21] Dietterich, T. (1995). Overfitting and undercomputing in machine learning. ACM computing surveys (CSUR), 27(3), 326-327. doi: 10.1145/212094.212114

[22] Ortiz, N., Hernández, R. D., Jimenez, R., Mauledeoux, M., & Avilés, O. (2018). Survey of biometric pattern recognition via machine learning techniques. Contemp. Eng. Sci., 11(34), 1677-1694. doi: 10.12988/ces.2018.84166 --

[23] Arjomandi, L., Khadka, G., Xiong, Z., & Karmakar, N. (2018). Document Verification: A Cloud-Based Computing Pattern Recognition Approach to Chipless RFID. IEEE Access, 6, 78007-78015. doi: 10.1109/access.2018.2884651 --