Journal of Intelligent System and Applied Data Science (JISADS)

Journal homepage : https://www.jisads.com
ISSN (2974-9840) Online

Machine Learning Techniques for Resource Management: A Survey Study

Hadiel Elshaik[1]* , *Salahe*ldin Edam[1]

hdola1989rm@gmail.com

[1]School of Electronic Engineering, Sudan University of Science and Technology, Khartoum, Sudan

Abstract

The study's objective was to use machine learning techniques to provide an overview of resource management issues. In order to demonstrate how resource management machine learning algorithm's function, the study uses a qualitative methodology. The results showed that the efficient deployment of very heterogeneous IoT networks depends on resource management, among other aspects of network administration. In conclusion, IoT networks struggle with resource allocation in addition to having to decide how to manage resources based on various contexts and situations. ML and DL models, unlike traditional resource management techniques such as optimization and heuristics-based methods, game theoretical and cooperative approaches, can derive actions from run-time context information and can retune and re-train themselves in response to changes in the environment. ML and DL approaches offer significant potential for managing and making decisions in IoT applications that are large-scale, complex, distributed, and dynamic. These approaches are particularly promising in addressing challenges related to model uncertainty, interpretability, training costs, and generalization from test workloads to real-world user workloads. To effectively tackle radio resource management issues in the expanding IoT networks, it is crucial to carefully design solutions and conduct further scientific research in the future.

Keywords*:* Machine Learning (ML); Resource Management; virtual machines (VMs); Deep Learning (DL); Reinforcement Learning (RL); Artificial Intelligence (AI); Heterogeneous Networks (HetNets)

1. INTRODUCTION

A range of resource management functions are handled by machine learning, including workload estimation, task scheduling, VM consolidation, resource optimization, and energy optimization (Khan et al., 2022). Computing's emergence as a fifth utility is presently underway possible as a result of the environment that cloud computing has created for consumers of software and IT infrastructure (Buyya et al., 2018). In cloud computing, data center resource management remains a tough problem that is significantly influenced by application workload. Traditional cloud computing infrastructures, such as data centers, where applications were connected to individual physical servers, were frequently over-provisioned in order to manage issues with the highest workload (Xu et al., 2017). Because of the waste of resources and floor space, the data center was costly to run in terms of resource management. On the other hand, virtualization technology has demonstrated its ability to make data centers easier to administer. Among the several benefits of this technology are server consolidation and increased server utilization. Large-scale technology giants like Amazon, Google, and Microsoft operate extensive data centres that require sophisticated resource management. These data centres encompass various components, including servers, virtual machines (VMs), and other associated management responsibilities (Bianchini et al., 2020). Many VMs with varying workload types and quantities are assigned to a server host in these data centers. Because of the variable and irregular demand, a server may be over- and underutilized, resulting in an imbalance in the resource use assigned to virtual machines on a certain hosting server. This could lead to issues such as inconsistent quality of service (QoS), unbalanced energy use, and SLA violations (Singh and Kumar, 2019).

In an uneven workload scenario, the average CPU and memory utilization was found to be 17.76% and 77.93%, respectively. However, studies conducted in Google data centres have shown that the CPU utilization in a Google cluster does not exceed 60%, while the memory utilization remains below 50% (Kumar et al., 2021). Workload inconsistency diminishes data center productivity, which has an impact on energy consumption. It is proportionate to the data centre's financial loss and operational expenses. Excess energy consumption impacts carbon footprints directly, and we must look for alternative and eliminate it because an ideal machine absorbs more than half of maximum energy usage. (Barroso et al., 2013). According to an EIA (Energy Information Administration) survey, data centres consumed around 35 Twh (Tera Watt hour) of energy in 2015 and will consume 95 Twh by 2040 (Khan et al., 2022). Determining the optimal mapping of virtual machines (VMs) to servers is crucial for balancing resource utilization and reducing the number of active servers (Li et al., 2013). However, this problem is challenging and falls under the category of NP-complete. To ensure quality of service (QoS) standards and maximize the benefits of data centres, it is essential to have an effective resource management strategy (Kumar and Singh, 2020). Intelligent mechanisms in the future will provide insights that enable applications to map to machines with higher resource utilization (Kumar et al., 2020). Predicting these future insights is challenging due to the nonlinear and dynamic nature of VM workloads.

Nevertheless, there are two methods for obtaining future workload insights: historical workload-based prediction methods, which learn trends from historical workload data, and homeostatic-based prediction methods, which estimate future workload by subtracting the prior workload from the current workload (Kumar and Singh, 2018). The mean of the prior workload can be either static or dynamic. Both approaches have advantages and disadvantages, but historical forecasts are considered more straightforward and well-established in this field (Khan et al., 2022).

Intelligent resource management will be crucial in maximizing the data center's SLA, energy consumption, and operational expenses by performing efficient and intelligent resource provisioning. Data center resource management includes tasks such as resource provisioning, reporting, workload scheduling, and a range of other responsibilities (Ilager et al., 2020). The provisioning of resources is central to many of these procedures. The purpose of resource provisioning is to provide cloud resources to virtual machines (VMs) in response to end-user requests while limiting SLA violations related to availability, dependability, response time constraints, and cost limits (Shahidinejad et al., 2021).

To minimize over- or under-provisioning, it should assign resources based on end-user requirements, such as allocating more or less resources to VMs. This resource allocation approach can be used in two ways: proactive and reactive.

Proactive tactics focus on anticipating future workloads by using historical workload trends as a guide, whereas reactive operations are carried out when resource demand emerges. As a result, It may be concluded that the experience of historical-based prediction methods can be successfully merged into proactive methods to provide intelligent dynamic resource scaling, hence promoting intelligent dynamic resource management. Furthermore, based on projections, different operations such as task scheduling, thermal management, and VM consolidation may be carried out to improve QoS and optimize resource utilization and energy usage. Machine learning (ML) techniques are used in a variety of industries, including computer vision, pattern recognition, and bioinformatics. The advancement of machine learning techniques has benefited large-scale computing systems (Mao et al., 2019). In a recent report, Google outlined its initiatives to optimize electricity use, reduce costs, and boost productivity (Jeff, 2018). The Structure of the research paper shown in Fig 1.

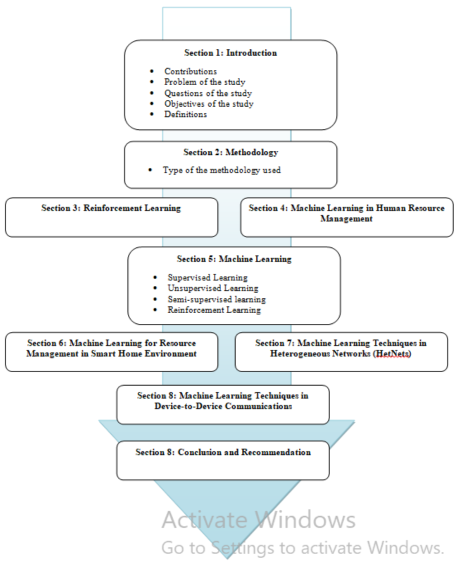

Figure1: Structure of the research paper.

2. REVIEW OF MACHINE LEARNING

Machine learning is used in resource management in a variety of ways, including workload estimation, job scheduling, VM consolidation, resource optimization, and energy optimization (Khan et al., 2022). The application burden has a substantial impact on data center resource management, which is still a complex problem. In conventional cloud computing environments such as data centers, applications were often connected to specific physical servers, and these servers were frequently over-provisioned to manage difficulties with the highest workload (Xu et al., 2017). Major IT behemoths like Google, Microsoft, and Amazon have enormous data centres with complex resource management. Servers, virtual machines (VMs), and other administrative duties are part of these massive data centres' resource management (Bianchini et al., 2020). In these data centres, a server host is assigned to a large number of VMs with varying workload types and amounts. The dynamic and fluctuating demand for resources in virtual machines can cause an imbalance in resource allocation on hosting servers, leading to over-utilization and under-utilization. These problems can result in irregular quality service (QoS), imbalanced energy utilization, and violations of service level agreements (SLAs) (Singh and Kumar, 2019). This paper aims to review the challenges associated with resource management and explores the application of machine learning techniques to address these issues.

Definitions

Machine Learning (ML) can be defined as the capacity to extrapolate knowledge from data and then apply that knowledge to modify the behavior of an ML agent in accordance with the learned information. Techniques for machine learning have been applied to classification, regression, and density estimation applications. IoT devices provide enormous amounts of data, which can be used by data-driven ML approaches to create automated IoT service solutions. Deep Learning (DL), more particularly machine learning (ML), can be utilized for feature extraction and practical categorization when there is a large and multidimensional amount of data accessible (Hussain et al., 2020).

3. METHODOLOGY

The goal of qualitative research, which was used in this study, is to find notable patterns that are indicative of a specific event through text analysis and interpretation, interviews, and observations (Auerbach and Silverstein, 2003). The steps of ML of wireless sensor networks shown in Fig. 2 [1].

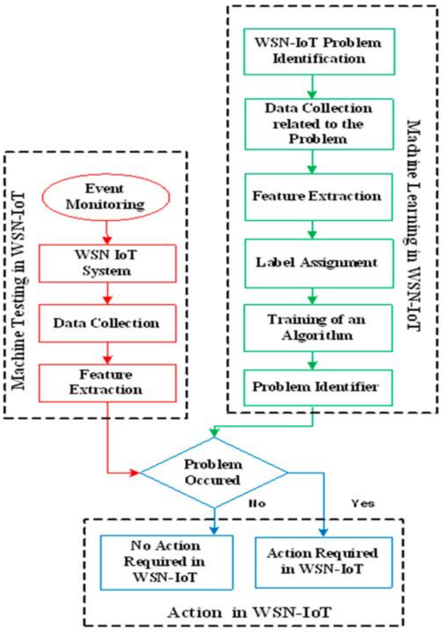

Figure 2: Steps of ML of wireless sensor networks.

Because it focuses on determining the characteristics of the population or subject under investigation, the study employs qualitative research as a research strategy. This qualitative methodology focuses on the "what" rather than the "why" of the study topic. The qualitative research approach focuses on describing the core of a demographic segment rather than "why" a certain occurrence happens. In other words, it "describes" the subject of the inquiry without going into detail about "why" it occurs (Maxwell, 2008).

The study followed the qualitative approach to explain the study (Machine Learning Techniques For Resource Management: A Survey Study) through the literature review.

Reinforcement Learning

We can apply ML approaches when past information about the system, network, users, and parameters is unavailable and needs to be anticipated along with control decisions. Reinforcement Learning (RL) is one of these methods, which involves monitoring system behavior and unknown parameters over time through trial and error in order to determine the best course of action. ML is advised when there is a model or algorithm lack for resource management challenges (Chen et al., 2019).

Model deficit refers to a lack of domain expertise or the absence of mathematical models, whereas algorithm deficit refers to the presence of a well-established mathematical model but difficulty optimizing existing algorithms using it. In this case, lower-complexity ML solutions are desirable. Furthermore, when contextual information is critical to include in the decision-making process, ML techniques are best suited. Because of the enormous number of devices producing massive amounts of data and the unknown system or network states and parameter values, the majority of IoT applications meet the aforementioned characteristics (Hussain et al., 2020).

4. MACHINE LEARNING IN HUMAN RESOURCE MANAGEMENT

Machine learning models are actively progressing in a variety of human resource management roles (Scholz, 2017). Currently, machine learning models are progressing in a number of areas related to human resource management. This study gives a summary of the important HR functions that can be enhanced by the implementation of machine learning and AI-based solutions. In this paper, three unique conceivable and potential scopes of AI solution implementation are examined, with a focus on three different aspects of employee engagement, organizational culture management, and the appraisal system. Using the decision tree model and logistic regression models to train datasets for an application might enhance the likelihood that the answers will be more accurate and will produce the best form of the evaluation system. If solutions are developed along the lines of what has been discussed, they may be helpful to organizations in managing their strategic human resource practices (Rudra Kumar, 2022).

Machine learning is the study of teaching computers to recognize objects or make predictions without being specifically programmed to do so (Jordan, 2015). Its primary concept is that by using statistical methods and training data, it is feasible to create algorithms that can forecast potential, unforeseen values. Over the previous two decades, machine learning has advanced from a research project to a widely utilized commercial tool. Machine learning has emerged as the go-to technique for creating useful applications in computer vision (Janai et al., 2020), speech recognition (Deng and Li, 2013), natural language processing (Olsson, 2009), robot control (Chin et al., 2020), self-driving cars (Stilgoe, 2018), efficient web search (Bhatia and Kumar, 2008), purchase recommendations (Hastie et al., 2009), and other artificial intelligence fields as shown in Fig. 3.[2].

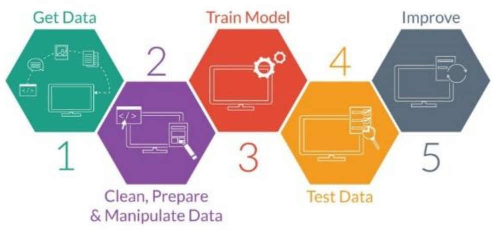

Figure 3: Machine Learning Process

Many AI system developers now understand that training a system by providing examples of acceptable input-output activities is significantly easier than manually programming it by making predictions for all possible inputs for various purposes. This accomplishment is primarily due to the availability of massive amounts of data as well as better server and GPU processing power efficiency [Goodfellow rt al., 2016]. Depending on the modeling aim and the issue at hand (RL), machine-learning algorithms are classified as supervised learning, semi-supervised learning (SSL), unsupervised learning, and reinforcement learning. Unsupervised learning is divided into two categories: clustering and dimension reduction (Hartigan et al., 1979; Guha et al., 2000; Ding et al., 2002), whereas supervised learning, is divided into two categories: regression problem and classification problem (e.g., sentence classification (Yoon, 2014; Wenpeng and Schütze, 2015), picture classification (Yang et al, 2009; Bazi and Melgani, 2009; Ciregan et al., 2013), etc.

Supervised Learning (Sen et al., 2020): In supervised learning, each data sample consists of a name and multiple input attributes. The objective of the learning process is to create a mapping function that accurately relates the input features to the corresponding label. This mapping function can then be used to predict the label for new data by utilizing additional input features. Supervised learning is a widely used machine learning approach across various applications. An example of supervised learning is classification, where an object is assigned to a specific category based on its characteristics, such as classifying a mobile device based on its brand and features. On the other hand, if the objective is to predict a continuous variable, such as stock prices, the supervised learning task is called regression.

Unsupervised Learning (Celebi and Aydin, 2016): When we have input features without corresponding labels, we are involved in unsupervised learning instead of supervised learning. Unsupervised learning aims to understand the underlying data distribution and explore patterns or differences among data points. A well-known example of unsupervised learning is the clustering problem, which aims to identify meaningful groups within the data, such as grouping virtual machines based on resource utilization patterns. Figure 4 [2] illustrates different machine learning methods, including unsupervised learning.

Semi-supervised learning [Van Engelen and Hoos, 2020]: This subfield of machine learning aims to combine these two tasks, leveraging information from one task to improve the performance of the other. Semi-supervised learning (SSL) algorithms often utilize unlabelled data points to enhance the classification process, for example, by utilizing additional data points with unknown labels. On the other hand, understanding the similarity or belongingness of certain data points to the same class can assist in guiding the clustering process. By incorporating labelled and unlabelled data, SSL approaches can benefit from the advantages of supervised and unsupervised learning to improve overall learning and inference capabilities.

Kober et al. (2013) define Reinforcement Learning as: RL differs from both supervised and unsupervised learning in several aspects. Unlike supervised learning, RL does not require labelled input/output pairs or explicit correction of inferior choices during training. Instead, RL involves an agent interacting with an environment, learning to make decisions through a balance between exploration and exploitation. The agent receives feedback in the form of rewards or penalties based on its actions, which guide its learning process to optimize long-term cumulative rewards. RL is particularly suitable for problems where an agent learns through trial and error to achieve a specific goal in dynamic and uncertain environments. The translator pays the agent for making wise choices or acting in a certain way. If not, it would be approved. Robotics and computer game agent science frequently employ reinforcement learning.

Figure 4: Types of machine learning

5. MACHINE LEARNING FOR RESOURCE MANAGEMENT IN SMART HOME ENVIRONMENT

Smart home applications are highly popular in the realm of IoT, as they integrate various technologies such as security cameras, handheld scanners, tablets, smart appliances, and wireless sensors. These devices often have diverse access and quality-of-service (QoS) requirements, and they access network resources in a random manner. To address resource allocation and random-access challenges within the smart home environment, ML methods like Q-learning and multi-armed bandit can prove beneficial. These techniques enable effective management and optimization of resources in smart homes. This is so that various methods of reinforcement learning may change with the network environment and learn to do so dynamically (Ali et al., 2020; Kanoun et al., 2016.).

Additionally, small sensors and small payload data can be grouped using K-means clustering and PCA, respectively. Be aware that running conventional optimization and heuristics-based techniques on small, low-cost, and energy-constrained sensors can be quite expensive computationally. Furthermore, because of the heterogeneity of the devices and the overhead associated with information updates and exchange, traditional game theoretical approaches might not be appropriate (Hussain, 2020).

6. MACHINE LEARNING TECHNIQUES IN HETEROGENEOUS NETWORKS (HETNETS)

Resource management in HetNets, where cellular and small cell users coexist with different Radio Access Technologies (RATs), poses several challenges, such as cross- and co-tier interference, mobility management, user association, RAT selection, and self-organization. To address these difficulties, researchers have explored ML-based methods alongside traditional techniques like optimization and heuristics (Omar et al., 2017). For instance, Simsek et al. (2015) developed an RL-based mechanism for inter-cell coordination and handover in HetNets, enabling devices to learn effective resource management. Vasudeva et al. (2017) utilized fuzzy game-theory techniques to reduce network energy consumption while maintaining QoS levels. Perez et al. (2017) proposed a unique cognitive RAT selection paradigm using ML methods.

Furthermore, ML approaches have also been employed for network self-organization, encompassing self-configuration, self-organization, and self-healing. Fan and Sengul (2014) investigated the use of artificial neural networks (ANN) for self-optimization in HetNets, while Alqerm and Shihada (2018) proposed an efficient resource allocation system utilizing online learning algorithms and Q-value theory for QoS provisioning at high data rates. A similar study on resource distribution in heterogeneous cognitive radio networks can be found in Fan and Sengul (2014).

7. MACHINE LEARNING TECHNIQUES IN DEVICE-TO-DEVICE COMMUNICATIONS

D2D networks allow two devices in close proximity to connect to one another without the need for a centralized base station. A D2D network offloads traffic from the primary BS by utilizing proximity communications, boosting the network's spectrum efficiency and Energy Efficiency (EE) (Ansari et al., 2017). Low route loss allows for high spectrum efficiency and sum rate, while low transmission power between radios guarantees EE. Numerous D2D network-related topics, including resource and power allocation, mode selection, proximity sensing, and interference avoidance, have been treated in the literature. (Ahmed et al., 2018; Liu et al., 2019). Recently, machine learning (ML) has been used to handle a range of D2D communication difficulties, including caching (Cheng et al., 2018), security and privacy [Haus et al., 2017], and others.

The efficient use of scarce resources to meet the QoS requirements of all network entities, including cellular and D2D users, presents a significant problem for D2D networks. The study of Maghsudi and Stańczak, (2014) developed a bandit-based channel access method for a distributed D2D system in which each pair chooses the best channel for communication. This raises the rates of individual D2D pairs while simultaneously reducing interference from other users sharing the same channels. Similarly to that, Asheralieva and Miyanaga, (2016) presented another study on channel selection with autonomous learning. The best resource selection technique is identified using Q-learning in a method for resource allocation in D2D networks that was published by Luo et al., (2014). Similar to Khan et al., (2017), the authors used cooperative RL to increase the individual device throughputs and the sum rates of the system by employing cooperative strategy planning to allocate resources optimally. In order to create an energy-efficient network solution, ML was used to optimize power allocations for various D2D couples (AlQerm and Shihada, 2017).

8. CONCLUSION

The aim of the study is to explain a review of problems with resource management utilizing machine learning techniques. The study employs a qualitative methodology to illustrate how resource management machine learning algorithms work. We employ the qualitative method to identify notable patterns that are suggestive of a specific occurrence.

In the context of IoT networks, resource management plays a vital role alongside other network administration tasks. The successful implementation of diverse IoT networks requires effective resource allocation and contextually appropriate resource management decisions. Unlike traditional approaches based on optimization and heuristics, game theoretical and cooperative methods are being utilized. ML and DL models, on the other hand, have the capability to adapt and retrain themselves by inferring actions from real-time context information in response to environmental changes. Particularly in complex, large-scale, distributed, and dynamic IoT application scenarios, ML and DL approaches hold significant promise for automating resource management and decision-making processes.

To solve complex radio resource management problems in emerging IoT networks, we recommend future scientific research. We must carefully build solutions for these networks in response to issues such as model uncertainty, model interpretability, model training costs, and generalization from test workloads to real application user workloads.

9. REFERENCES

[1] Asheralieva, A., & Miyanaga, Y. (2016). An autonomous learning-based algorithm for joint channel and power level selection by D2D pairs in heterogeneous cellular networks. IEEE transactions on communications, 64(9), 3996-4012.

[2] Cheng, P., Ma, C., Ding, M., Hu, Y., Lin, Z., Li, Y., & Vucetic, B. (2018). Localized small cell caching: A machine learning approach based on rating data. IEEE Transactions on Communications, 67(2), 1663-1676.

[3] Fang, Z. (2010). Resource management on cloud systems with machine learning (Master's thesis, Universitat Politècnica de Catalunya).

[4] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[5] Ali, S., Ferdowsi, A., Saad, W., Rajatheva, N., & Haapola, J. (2020). Sleeping multi-armed bandit learning for fast uplink grant allocation in machine type communications. IEEE Transactions on Communications, 68(8), 5072-5086.

[6] Ding, C., He, X., Zha, H., & Simon, H. D. (2002, December). Adaptive dimension reduction for clustering high dimensional data. In 2002 IEEE International Conference on Data Mining, 2002. Proceedings. (pp. 147-154). IEEE.

[7] Auerbach, C., & Silverstein, L. B. (2003). Qualitative data: An introduction to coding and analysis (Vol. 21). NYU press.

[8] Bhatia, M. P. S., & Kumar, A. (2008). Information retrieval and machine learning: supporting technologies for web mining

research and practice. Webology, 5(2), 5.

[9] Bazi, Y., & Melgani, F. (2009). Gaussian process approach to remote sensing image classification. IEEE transactions on geoscience and remote sensing, 48(1), 186-197.

[10] Bianchini, R., Fontoura, M., Cortez, E., Bonde, A., Muzio, A., Constantin, A. M., ... & Russinovich, M. (2020). Toward ml-centric cloud platforms. Communications of the ACM, 63(2), 50-59.

[11] Buyya, R., Srirama, S. N., Casale, G., Calheiros, R., Simmhan, Y., Varghese, B., ... & Shen, H. (2018). A manifesto for future generation cloud computing: Research directions for the next decade. ACM computing surveys (CSUR), 51(5), 1-38.

[12] Ciregan, D., Meier, U., & Schmidhuber, J. (2012, June). Multi-column deep neural networks for image classification. In 2012 IEEE conference on computer vision and pattern recognition (pp. 3642-3649). IEEE.

[13] Chen, M., Challita, U., Saad, W., Yin, C., & Debbah, M. (2019). Artificial neural networks-based machine learning for wireless networks: A tutorial. IEEE Communications Surveys & Tutorials, 21(4), 3039-3071.

[14] Ahmed, M., Li, Y., Waqas, M., Sheraz, M., Jin, D., & Han, Z. (2018). A survey on socially aware device-to-device communications. IEEE Communications Surveys & Tutorials, 20(3), 2169-2197.

[15] Barroso, L. A., Clidaras, J., & Hölzle, U. (2013). The datacenter as a computer: An introduction to the design of warehouse-scale machines. Synthesis lectures on computer architecture, 8(3), 1-154.

[16] Guha, S., Rastogi, R., & Shim, K. (2000). ROCK: A robust clustering algorithm for categorical attributes. Information systems, 25(5), 345-366.

[17] Celebi, M. E., & Aydin, K. (Eds.). (2016). Unsupervised learning algorithms. Berlin: Springer International Publishing.

[18] Deng, L., & Li, X. (2013). Machine learning paradigms for speech recognition: An overview. IEEE Transactions on Audio, Speech, and Language Processing, 21(5), 1060-1089.

[19] Ansari, R. I., Chrysostomou, C., Hassan, S. A., Guizani, M., Mumtaz, S., Rodriguez, J., & Rodrigues, J. J. (2017). 5G D2D networks: Techniques, challenges, and future prospects. IEEE Systems Journal, 12(4), 3970-3984.

[20] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. Journal of the royal statistical society. series c (applied statistics), 28(1), 100-108.

[21] Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

[22] Haus, M., Waqas, M., Ding, A. Y., Li, Y., Tarkoma, S., & Ott, J. (2017). Security and privacy in device-to-device (D2D) communication: A review. IEEE Communications Surveys & Tutorials, 19(2), 1054-1079.

[23] Hussain, F., Hassan, S. A., Hussain, R., & Hossain, E. (2020). Machine learning for resource management in cellular and IoT networks: Potentials, current solutions, and open challenges. IEEE communications surveys & tutorials, 22(2), 1251-1275.

[24] Ilager, S., Muralidhar, R., & Buyya, R. (2020, October). Artificial intelligence (ai)-centric management of resources in modern distributed computing systems. In 2020 IEEE Cloud Summit (pp. 1-10). IEEE.

[25] Sen, P. C., Hajra, M., & Ghosh, M. (2020). Supervised classification algorithms in machine learning: A survey and review. In Emerging technology in modeling and graphics (pp. 99-111). Springer, Singapore.

[26] Kumar, J., Singh, A. K., & Buyya, R. (2021). Self-directed learning-based workload forecasting model for cloud resource management. Information Sciences, 543, 345-366.

[27] Stilgoe, J. (2018). Machine learning, social learning and the governance of self-driving cars. Social studies of science, 48(1), 25-56.

[28] Xu, M., Tian, W., & Buyya, R. (2017). A survey on load balancing algorithms for virtual machines placement in cloud computing. Concurrency and Computation: Practice and Experience, 29(12), e4123.

[29] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

[30] Scholz, T. M. (2017). Big data in organizations and the role of human resource management: A complex systems theory-based conceptualization. Frankfurt a. M.: Peter Lang International Academic Publishers.

[31] Kumar, J., & Singh, A. K. (2020). Cloud datacenter workload estimation using error preventive time series forecasting models. Cluster Computing, 23(2), 1363-1379.

[32] Liu, S., Wu, Y., Li, L., Liu, X., & Xu, W. (2019). A two-stage energy-efficient approach for joint power control and channel allocation in D2D communication. IEEE Access, 7, 16940-16951.

[33] Maghsudi, S., & Stańczak, S. (2014). Channel selection for network-assisted D2D communication via no-regret bandit learning with calibrated forecasting. IEEE Transactions on Wireless Communications, 14(3), 1309-1322.

[34] Wenpeng, Y., & Schütze, H. (2015). Multichannel variable-size convolution for sentence classification [C]. In Proc of the 19th Conf on Computational Natural Language Learning. Stroudsburg, PA: ACL (pp. 204-214).

[35] Janai, J., Güney, F., Behl, A., & Geiger, A. (2020). Computer vision for autonomous vehicles: Problems, datasets and state of the art. Foundations and Trends® in Computer Graphics and Vision, 12(1–3), 1-308.

[36] Kumar, J., Singh, A. K., & Buyya, R. (2020). Ensemble learning based predictive framework for virtual machine resource request prediction. Neurocomputing, 397, 20-30.

[37] Yoon, K. (2014). Convolutional Neural Networks for Sentence Classification [OL]. arXiv Preprint.

[38] Mao, H., Schwarzkopf, M., Venkatakrishnan, S. B., Meng, Z., & Alizadeh, M. (2019). Learning scheduling algorithms for data processing clusters. In Proceedings of the ACM special interest group on data communication (pp. 270-288).

[39] Khan, T., Tian, W., Zhou, G., Ilager, S., Gong, M., & Buyya, R. (2022). Machine learning (ML)–Centric resource management in cloud computing: A review and future directions. Journal of Network and Computer Applications, 103405.

[40] Simsek, M., Bennis, M., & Güvenç, I. (2015, March). Context-aware mobility management in HetNets: A reinforcement learning approach. In 2015 ieee wireless communications and networking conference (wcnc) (pp. 1536-1541). IEEE.

[41] Van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. Machine Learning, 109(2), 373-440.

[42] Perez, J. S., Jayaweera, S. K., & Lane, S. (2017, June). Machine learning aided cognitive RAT selection for 5G heterogeneous networks. In 2017 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom) (pp. 1-5). IEEE.

[43] Yang, J., Yu, K., Gong, Y., & Huang, T. (2009, June). Linear spatial pyramid matching using sparse coding for image classification. In 2009 IEEE Conference on computer vision and pattern recognition (pp. 1794-1801). IEEE.

[44] Omar, M. S., Hassan, S. A., Pervaiz, H., Ni, Q., Musavian, L., Mumtaz, S., & Dobre, O. A. (2017). Multiobjective optimization in 5G hybrid networks. IEEE Internet of Things Journal, 5(3), 1588-1597.

[45] Shahidinejad, A., Ghobaei-Arani, M., & Masdari, M. (2021). Resource provisioning using workload clustering in cloud computing environment: a hybrid approach. Cluster Computing, 24(1), 319-342.

[46] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

[47] Li, X., Qian, Z., Lu, S., & Wu, J. (2013). Energy efficient virtual machine placement algorithm with balanced and improved resource utilization in a data center. Mathematical and Computer Modelling, 58(5-6), 1222-1235.

[48] Jeff, D. (2018). ML for system, system for ML, keynote talk in Workshop on ML for Systems, NIPS.

[49] Luo, Y., Shi, Z., Zhou, X. I. N., Liu, Q., & Yi,

Q. (2014, December). Dynamic resource allocations based on Q-learning for D2D communication in cellular networks. In 2014 11th international computer conference on wavelet active media technology and information processing (ICCWAMTIP) (pp. 385-388). IEEE.

[50] Kanoun, K., Tekin, C., Atienza, D., & Van Der Schaar, M. (2016). Big-data streaming applications scheduling based on staged multi-armed bandits. IEEE Transactions on Computers, 65(12), 3591-3605.

[51] Kumar, J., & Singh, A. K. (2018). Workload prediction in cloud using artificial neural network and adaptive differential evolution. Future Generation Computer Systems, 81, 41-52.

[52] Simsek, M., Bennis, M., & Güvenç, I. (2015, March). Context-aware mobility management in HetNets: A reinforcement learning approach. In 2015 IEEE Wireless Communications and Networking Conference (WCNC) (pp. 1536-1541). IEEE.