VOL 4 (2020) NO 3

e-ISSN : 2549-9904
ISSN : 2549-9610

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

Reinforcement Learning Rebirth, Techniques, Challenges, and Resolutions

Wasswa Shafik#, S. Mojtaba Matinkhah#, Parisa Etemadinejad#, Mamman Nur Sanda*

Computer Engineering Department, Yazd University, Yazd, Iran
E-mail: wasswashafik@stu.yazd.ac.ir, matinkhah@yazd.ac.ir, Parisa.etemadinejad@stu.yazd.ac.ir

* Department of Physics, Yazd University, Yazd, Iran
E-mail: mammannur10@stu.yazd.ac.ir

***Abstract*— Reinforcement learning (RL) is a new propitious research space that is well-known nowadays on the internet of things (IoT), media and social sensing computing are addressing a broad and pertinent task through making decisions sequentially by deterministic and stochastic evolutions. The IoTs extend world connectivity to physical devices like electronic devices network by use interconnect with others over the Internet with the possibility of remotely being supervised and meticulous. In this paper, we comprehensively survey an in-depth assessment of RL techniques in IoT systems focusing on the main known RL techniques like artificial neural network (ANN), Q-learning, Markov Decision Process (MDP), Learning Automata (LA). This study examines and analyses learning technique with focusing on challenges, models performance, similarities and the differences in IoTs accomplish with most correlated proposed state of the art models. The results obtained can be used as a foundation for designing, a model implementation based on the bottlenecks currently assessed with an evaluation of the most fashionable hands-on utility of current methods for reinforcement learning.**

***Keywords*— internet of things; reinforcement learning; ANN; learning automata; q-learning; markov decision process.**

I. INTRODUCTION

Reinforcement Learning (RL) is a category of the Machine Learning (ML) techniques that are decided, supervised, semi- supervised and unsupervised besides that is also a division of Artificial Intelligence (AI). It consents machines and software agents to automatically determine the ideal behavior within a specific context, with an attempt to maximize performance [1]. Unassertive return is vital for the agent to learn its behavior is referred to as the reinforcement signal. Social media platforms that embedded on IoT devices utilize RL for instance automatically tags people and identify common objects like landmarks in uploaded pictures among more. Different numbers of the algorithm that tackle this applicability and automatic recovery of data are considered with the time of learning and are now available [2].

RL examines and evaluates a detailed sort of problem, with all its resolutions are referred to as RL algorithms. It is applied in many categories of technology phenomena like detecting the premature onset of an infection, fraud detection, resource optimization, programmed or self-driving cars, facial recognition, high volume trading among more with real-valued function [3]. Computing categorized as dynamic programming that trains algorithms by means of a system of return and penalty. The learning holds some studying patterns to the approach of data detection including categorization, prediction, and identification. This kind of automated learning scheme indicates that there is little requirement for a human expert who knows about the domain [4].

RL is challenged with memory extensively to store values of each state, since the problems are a times complex, solving this involves observing value approximation techniques, like neural networks [5]. There are many connotations of introducing these imperfect value estimations and research tries to minimize their influence on the quality and the authentication enhancement where IoT is managed and maintained using this RL entity as illustrated in figure1 depicting approaches, challenges with applications. We observed that it is a sign to carry out the survey and provide the researcher with a piece of summarized information about for RL in IoT so that in case there is a need to create algorithms and models, there is an easy approach. We mainly categorized RL techniques. This paper uniquely focuses on the following areas as summarized:

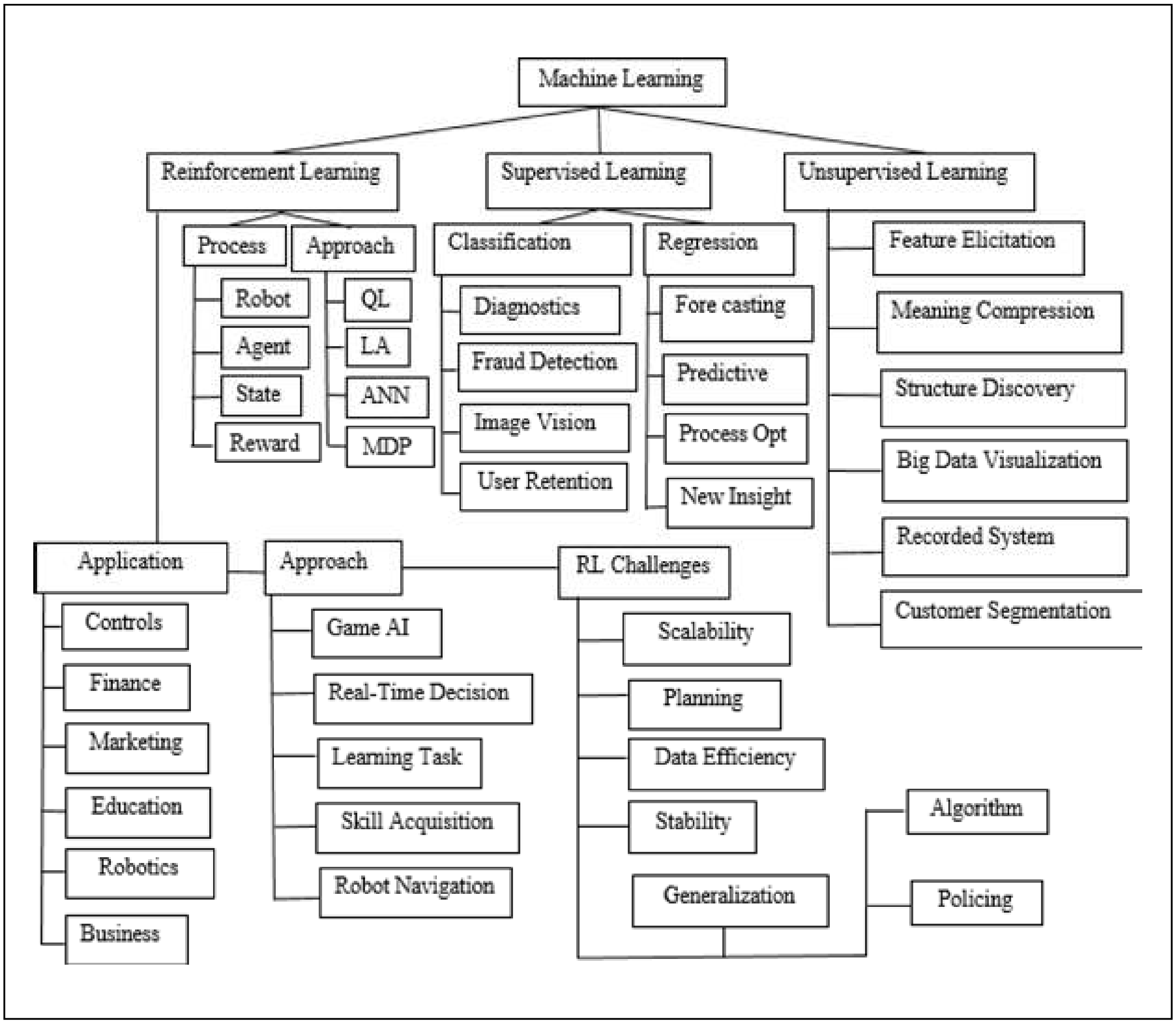

Fig. 1 Approaches, Challenges, and Applications of Reinforcement Learning

- Present a comprehensive and in-depth systematic survey of the main reinforcement learning techniques in IoTs.
- Describe current state-of-the-art results solution on IoT networks with a close focus on the reinforcement learning techniques in the IoTs.
- Examine and describe the relationship between IoTs and the reinforcement learning techniques based on application, issues, and resolutions.
- Present an in-depth review of existing studies solutions and models to the challenges identified related to IoT application and enlighten on internet connectivity.
- Afford summarized tables that categorize these reinforcement learning techniques in different phenomena that cut across in resemblance, identified independent challenges.

Due to the limited perception, regularly impossibilities to determine the current state is the problem in this area of research, this affects the performance of the set of rules. Issues like applicable rules might be intuited, but are not easily designated by unpretentious logical rules, potential outputs are defined but which action to take is dependent on diverse circumstances which cannot be predicted, accuracy is supplementary significant than interpretation or interpretability [6, 7].

The rest of this paper is structured as follows; in section 2 provides related work. In section 3, illuminates RL in general availing brief information about techniques. In section 4, explicates the classification of the RL in IoT in table 1 with state of the earth solutions in table 2. Finally, sections 5 have the conclusion the article and indistinguishably depict our future work

II. Related Work

In this section, we discuss different areas where RL techniques have been applied with the ability of the machines to practice and learning is recognized as algorithms.

Within the security phenomena and its associated challenges including attacks [8], confidentiality and integrity, physical access within the IoTs analysis on the standard and natural policy gradients on actor-critics [9], huge or big data processing in learning [10], user simulation techniques for RL example dialogue management strategies [11], robotic systems during learning, node discovery within IoTs scenarios [12], content-aware computing with a close focus on the learning and data screening analytics [13]. In appreciation to the existing works, they have not summarised the existing solutions to the challenges as this study avails in the summarized tables and figures involved.

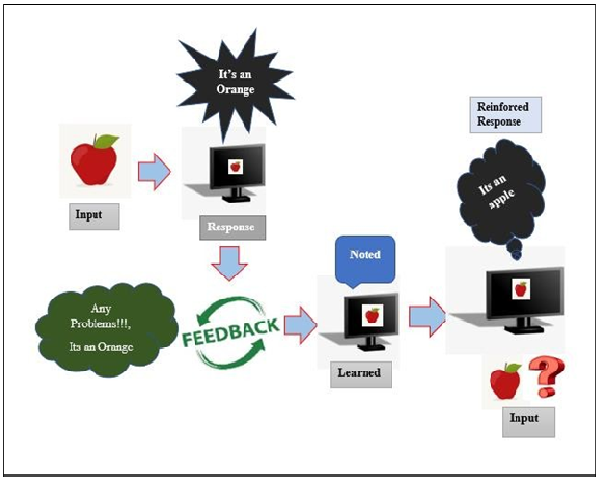

Fig. 2 Expressive Reinforcement Learning Procedure

Within the security phenomena and its associated challenges including attacks [8], confidentiality and integrity, physical access within the IoTs analysis on the standard and natural policy gradients on actor-critics [9], huge or big data processing in learning [10], user simulation techniques for RL example dialogue management strategies [11], robotic systems during learning, node discovery within IoTs scenarios [12], content-aware computing with a close focus on the learning and data screening analytics [13]. In appreciation to the existing works, they have not summarised the existing solutions to the challenges as this study avails in the summarized tables and figures involved.

III. Reinforcement Learning Techniques

In this section, RL techniques are presented and summarized in figure 2. The machine is provided with a set of acceptable actions, rules, and potential end states. By smearing the rules, exploring different actions and detecting resulting reactions the machine learns to adventure the rules to generate the desired result. Accordingly, determining what sequence of actions, in what surroundings, resolves to an optimized result. Mathematical algorithms and programming in space search, statistical and dynamic programming to estimate the utility of different learning aspects.

RL necessitated a lot of data, consequently, it is relevant in domains where simulated data is readily available identical to gameplay, robotics [14], [116]. Other areas include text mining or text summarization engines, dialogue agent trade transaction, health care, and navigations. Therefore, the four major techniques of RL is briefly explained below:

A. *Artificial* Neural Networks (ANN)

Neural networks are sometimes called connectionist systems that use computational algorithms and capable of pattern recognition. RL is accessible as systems of interconnected “neurons” which can compute values from inputs. It is based on a collection of connected nodes called artificial neurons that loosely model the neurons in a biological brain [15], [117].

ANN is currently used including feedforward neural network, radial basis function neural network, Recurrent Neural Network (RNN) Long Short-Term Memory, Convolutional neural networks, and Modular Neural Networks. Some of the advantages of these techniques include the ability to work with incomplete knowledge, fault tolerance, having a distributed memory, Parallel processing capability, ability to make machine learning [16].

B. *Learning* Automata

Early learning techniques that use adaptive decision-making with unit situated in a random environment that absorbs the optimal action over frequent relations.

The arrangements are selected according to an explicit probability distribution which is efficiently constructed on the situation response on the automation obtains by execution a specific accomplishment [17], [118]. LA managed a multipart, highly non-linear, indefinite and half-finished have to delicate and interactive exchange with the environment where they operate [18].

C. *Markov Decision Processes (MDP)*

MDP has an isolated time stochastic control procedure providing a mathematical framework for modelling verdict creation in situations where outcomes are partly random and partly under the control of a result maker. The resolution for an MDP is a policy that designates the superlative action for each state in the MDP called the optimal policy found through a variety of methods, like dynamic programming. The difference between LA and Q-learning (QL) is that the former technique neglects the memory of Q-values, but updates the action possibility straight to find the learning result. LA is a learning scheme with a rigorous proof of convergence [19], [119].

D. *Q-Learning*

The penalty area of QL is to absorb a policy, which expresses an agent pardon's action to take under what surroundings does not even necessitate a model of the environment and it can grip difficulties with stochastic transitions and plunders, deprived of necessitating adaptations [20]. The penalty area of QL is to absorb a policy, which expresses an agent pardons action to take under what surroundings that does not even necessitate a model of the environment and it can grip difficulties with stochastic transitions and plunders, deprived from necessitating adaptations [20], [120].

TABLE I
TECHNIQUES CLASSIFICATION BASED ON THE APPLICATION

Technique classification	Classification	
	IoT Issue-Based	**Application Based**
ANN	Intrusion prediction [23] IoT representation annotation [24] Data-driven management [25] Data and Feedback validation [26] Visualization and understanding [27] Learning environment detection [28] Fraud detection [29]	Prediction of the performance [50] Classification of capability [51] Tolerance related acquisition [52] IoT crime forensics [53] Fraud detection in IoT application [54] IoT decision process and making [55]
LA	Intrusion prediction [30] IoT representation annotation [31] Data-driven management [32] Data and Feedback validation [33] Visualization and understanding [34] Learning environment detection [35] Fraud detection [36]	Predicting Software Defects on IoTs [56] Prediction of behavioral changes [57] Signature verification [58] Analysis and decisions [59] Auto-selection of IoT task [60] Traffic incident detection [61] Telecommunication [62] Internet networks [63]
MDP	Intrusion prediction [37] IoT representation annotation [38] Data-driven management [39] Data and Feedback validation [40] Visualization and understanding [41] Learning environment detection [42] Fraud detection [43]	Reinforcement Recognition [64] Short-term traffic forecasting [65] long-term traffic flow forecasting [66] Face recognition [67] Speech and text recognition [68] Data classification [69]
QL	Intrusion prediction [44] IoT representation annotation [45] Data-driven management [46] Data and Feedback validation [47] Visualization and understanding [48] Learning environment detection [49]	IoT decision and processing division [70] IoT Induction detection [71] Navigational IoT detection [72] IoT fault diagnosis [73]

TABLE II
CLASSIFICATION BASED ON THE CURRENT STATE OF THE ART SOLUTIONS

Technique	**Some Identified issue**	**State of the art Solution**	**Reference**
ANN	Intrusion prediction IoT representation annotation Data-driven management Data and Feedback validation Visualization and understanding Learning environment detection Fraud detection Tolerance related acquisition	Precognitive ANN algorithm Hybrid NN for document classification Management models based on Biases A neural-fuzzy model Design Deep generating	[76] [77] [78] [79] [80] [81] [82] [83]
LA	Intrusion prediction IoT representation annotation Data-driven management Data and Feedback validation Visualization and understanding Learning environment detection Fraud detection	Development of the LA modes Wave font cellar LA Computation and data-driven modeling IPTV viewer modeling Probabilistic methodologies	[84] [85] [86] [87] [88] [89] [90]

MDP	Intrusion prediction IoT representation annotation Data-driven management Data and Feedback validation Visualization and understanding Learning environment detection Fraud detection	Filter models Code retrieval Multi-period decision-making models AI integration with Neurodegenerative A Self-supervised Approach	[91] [92] [93] [94] [95] [96] [97]
QL	Intrusion prediction IoT representation annotation Data-driven management Data and Feedback validation Visualization and understanding Learning environment detection Fraud detection	Deep computation model Distant supervision relation Extractor Fault data management ReNeg and backseat driver Human- level control	[98] [99] [100] [101] [102] [103] [104]

QL holds different variants including deep Q-learning, double Q-learning, delayed Q-learning and the greedy Q-learning used in the combination with function approximation and convergence is guaranteed even when function approximation is used to estimate the action values is an advantage [21].

A dynamic decision-making unit positioned in an arbitrary environment that acquires the optimal action through repetitive connections with its environment [74]. The activities are chosen to render specific probability circulation which is updated based on the environment response the automation attained by execution with a particular action. LA is presently applied in most irregular patterns including photo, snap, or image dispensation, graph complexion, social modeling, collecting and sensor network corresponding to the channel obligation routing among others [75].

IV. Classification of Reinforcement Learning

In this section, elementary applications, issues, and solutions for most current models are discussed. These types include positive reinforcement, negative reinforcement, punishment, and extinction. Below are some of the issues associated with RL techniques which are arranged according to the impact during the learning process.

A. *Reinforcement Learning Applicability*

An ANN is constructed for an explicit application, like pattern recognition or data classification, over a learning process. Learning largely involves adjustments to the synaptic connections that exist between the neurons with entities including Interconnections, learning rules [21]. ANN holds five basic categories of neuron connection that include a single-layer feed-forward network, a multilayer feed-forward network is a single node with its own feedback, a single-layer recurrent network, and lastly multilayer recurrent network [22]. In table 1, we present a summarized classification of the IoT aspects based on the IoT issues and application.

V. State of The Art Solution

Within this section, we presented some merits of RL in everyday activities such as holding a comprehensive conversation below in table 2, we provide the classification of the techniques based on the IoT issues. IoT applications and ANN in the smart world including smart houses, smart card and smart city among others. IoT is receiving countless attention due to its probable strength and ability to be integrated into any complex structures and it is becoming a great tool to acquire data from a particular environment to the cloud [105]. In smart transportation, today, covers route optimization, parking, street lights, accident anticipation/detection, road anomalies, and infrastructure IoT applications in Intelligent Transportation Systems (ITS) and obtain a clear view of the trends in the aforementioned fields and spot thinkable attention requests [106], [121].

A. *Physical Data Entry*

Availed erroneousness and duplication of data are major IoT organization- based underprovided to automate its processes [107], [115]. RL set of rules and extrapolative modeling algorithms can expressively improve this situation.

RL uses the exposed data to progress the process as more multiplication is made. Accordingly, now devices can acquire to accomplish time-intensive certification and data access responsibilities, familiarity workers can now devote more time on higher-value problem-solving responsibilities [108].

B. *Detecting Junks*

For instance, email capability providers used pre-existing rule-based presentations to remove junk. Nevertheless, now the junk filters create new rules themselves using RL were junk sometimes direct mail detection is the earliest problem solved by neural networks techniques in its junk filters [109]. It is noted that like Google now a day boasted the proportion of junk rates since now recognition of this junk mail and phishing messages by analyzing rules across an enormous collection of computers is possible [110].

C. Merchandise Approval

RL has permitted today a merchandise-based endorsement system since models can identify those products in which that purchaser drives be attentive and perspective to acquisitions. The RL algorithm recognizes hidden patterns amongst substances and emphases on an alliance of similar products into bands [111]. The RL model of this decision procedure would permit a program to brand approval to a purchaser and motivate product purchases sideways with section detail is used by social media to commendation users to connect with other operators [112].

VI. CONCLUSION

In this paper, various prevalent classification techniques of RL have been discussed with their elementary approaches to application, challenges, and state-of-the-solution. Classification procedures were based on the application, challenge, and state of the art solutions that are implemented be implemented on the different type of data sets like in IoT setups synchronization, resource optimization, consumption efficiency among more.The study discovered that all RL technique is much superior to it comes to IoT systems and usage since separate techniques hold their own compensations, downsides and execution issues [113,114]. The selection of classification techniques depends on user problematic field of approach to usage. This research provides an opening approach to challenges affecting RL in IoT and denoting unapproached solutions. Through this, we got interested in extending this study more deeply towards IoT systems by designing a model to handle the dynamics of the next wireless generation (fifth generation computation).

FUNDING

The research in this publication was supported under Intelligent Connectivity Research laboratory Project with full Research Grant from the Computer Engineering Department, Yazd University, Iran.

AVAILABILITY OF DATA AND MATERIALS

All materials and related literature to this survey research have been publicly included in this publication and exposed.

AUTHORS' CONTRIBUTIONS

All authors have participated in (a) conception and design, or analysis and interpretation of the literature; (b) drafting the article or revising it critically for important intellectual content, and (c) approval of the final version.

COMPETING INTERESTS

The authors declare that they have no competing interests totally.

ACKNOWLEDGMENT

We would like to extend appreciation and express our gratitude to the Department members of the Computer Engineering faculty for their remarks and assistance for us to accomplish the superiority of this paper.

REFERENCES

[1] J. Su, D. V. Vargas, and K. Sakurai, "Attacking convolutional neural network using differential evolution," IPSJ Trans. Comput. Vis. Appl., vol. 11, no. 1, p. 1, 2019.
[2] A. Sakata, N. Takemura, and Y. Yagi, "Gait- based age estimation using multi-stage convolutional neural network," IPSJ Trans. Comput. Vis. Appl., vol. 11, no. 1, p. 4, 2019.
[3] V. Gullapalli, "A stochastic reinforcement learning algorithm for learning real-valued functions," Neural networks, vol. 3, no. 6, pp. 671–692, 1990.
[4] W. D. Smart and L. P. Kaelbling, "Effective reinforcement learning for mobile robots," in Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292), 2002, vol. 4, pp. 3404–3410.
[5] J. Garcia, F. R. Ervin, and R. A. Koelling, "Learning with prolonged delay of reinforcement," Psychonomic Science, vol. 5, no. 3, pp. 121–122, 1966.
[6] L. K. Fellows and M. J. Farah, "Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm," Brain, vol. 126, no. 8, pp. 1830–1837, 2003.
[7] B. Chatterjee, D. Das, S. Maity, and S. Sen, "RF- PUF: Enhancing IoT Security through Authentication of Wireless Nodes using In-situ Machine Learning," IEEE Internet of Things Journal, vol. 6, no. 1, pp. 388–398, 2019.
[8] P. Sun, J. Li, M. Z. A. Bhuiyan, L. Wang, and B. Li, "Modeling and clustering attacker activities in IoT through machine learning techniques," Information Sciences, vol. 479, pp. 456–471, 2019.
[9] A. Singla and A. Sharma, "Physical Access System Security of IoT Devices using Machine Learning Techniques," Available at SSRN 3356785, 2019.
[10] P. Punithavathi, S. Geetha, M. Karuppiah, S. H. Islam, M. M. Hassan, and K.-K. R. Choo, "A lightweight machine learning-based authentication framework for smart IoT devices," Information Sciences, vol. 484, pp. 255–268, 2019.
[11] F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain, "Machine Learning in IoT Security: Current Solutions and Future Challenges," arXiv preprint arXiv:1904.05735, 2019.
[12] I. Grondman, L. Busoniu, G. A. Lopes, and R. Babuska, "A survey of actor-critic reinforcement learning: Standard and natural policy gradients," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no. 6, pp. 1291–1307, 2012.
[13] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," EURASIP Journal on Advances in Signal Processing, vol. 2016, no. 1, p. 67, 2016.
[14] J. Schatzmann, K. Weilhammer, M. Stuttle, and S. Young, "A survey of statistical user simulation techniques for reinforcement-learning of dialogue management strategies," The knowledge engineering review, vol. 21, no. 2, pp. 97–126, 2006.
[15] Z. Shi, J. Tu, Q. Zhang, L. Liu, and J. Wei, "A survey of swarm robotics system," in International Conference in Swarm Intelligence, 2012, pp. 564– 572.
[16] M. L. Valarmathi, L. Sumathi, and G. Deepika, "A survey on node discovery in Mobile Internet of Things (IoT) scenarios," in 2016 3rd International Conference on Advanced Computing and Communication Systems (ICACCS), 2016, vol. 1, pp. 1–5.
[17] O. B. Sezer, E. Dogdu, and A. M. Ozbayoglu, "Context-aware computing, learning, and big data in internet of things: a survey," IEEE Internet of Things Journal, vol. 5, no. 1, pp. 1–27, 2018.
[18] R. Ahad, and Y. Yagi, "Spatio-temporal silhouette sequence reconstruction for gait recognition against occlusion," IPSJ Trans. Comput. Vis. Appl., vol. 11, no. 1, p. 9, 2019.
[19] P. Goyal, H. Malik, and R. Sharma, "Application of Evolutionary Reinforcement Learning (ERL) Approach in Control Domain: A Review," in Smart Innovations in Communication and Computational Sciences, Springer, 2019, pp. 273–288.
[20] Z. Liu, C. Yao, H. Yu, and T. Wu, "Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things," Future Generation Computer Systems, 2019, pp 1-9.
[21] [21] M. Shojafar and M. Sookhak, Internet of everything, networks, applications, and computing systems (IoENACS). Taylor & Francis, 2019, pp 1-3.
[22] M. Jamshidi, S. S. A. Poor, N. N. Qader, M. Esnaashari, and M. R. Meybodi, "A Lightweight Algorithm against Replica Node Attack in Mobile Wireless Sensor Networks using Learning Agents," IEIE

Transactions on Smart Processing & Computing, vol. 8, no. 1, pp. 58–70, 2019.
[23] X. Lu, Y. Tsao, S. Matsuda, and C. Hori, "Speech enhancement based on deep denoising autoencoder.," in Interspeech, 2013, pp. 436–440.
[24] H. Leopold, H. van der Aa, J. Offenberg, and H. A. Reijers, "Using Hidden Markov Models for the accurate linguistic analysis of process model activity labels," Information Systems, vol. 83, pp. 30–39, 2019.
[25] T. Chen, S. Barbarossa, X. Wang, G. B. Giannakis, and Z.-L. Zhang, "Learning and Management for Internet of Things: Accounting for Adaptivity and Scalability," Proceedings of the IEEE, vol. 107, no. 4, pp. 778–796, 2019.
[26] S. A. M. Shihab, C. Logemann, D.-G. Thomas, and P. Wei, "Towards the Next Generation Airline Revenue Management: A Deep Reinforcement Learning Approach to Seat Inventory Control and Overbooking," arXiv preprint arXiv:1902.06824, 2019.
[27] R. Rocchetta, L. Bellani, M. Compare, E. Zio, and E. Patelli, "A reinforcement learning framework for optimal operation and maintenance of power grids," Applied Energy, vol. 241, pp. 291–301, 2019.
[28] R. Vafashoar and M. R. Meybodi, "Reinforcement learning in learning automata and cellular learning automata via multiple reinforcement signals," Knowledge-Based Systems, vol. 169, pp. 1–27, 2019.
[29] X. Qi, Y. Luo, G. Wu, K. Boriboonsomsin, and M. Barth, "Deep reinforcement learning enabled self- learning control for energy efficient driving," Transportation Research Part C: Emerging Technologies, vol. 99, pp. 67–81, 2019.
[30] F. Zantalis, G. Koulouras, S. Karabetsos, and D. Kandris, "A Review of Machine Learning and IoT in Smart Transportation," Future Internet, vol. 11, no. 4, p. 94, 2019.
[31] J. Muñuzuri, L. Onieva, P. Cortés, and J. Guadix, "Using IoT data and applications to improve port- based intermodal supply chains," Computers & Industrial Engineering, 2019.
[32] N. Jiang, Y. Deng, O. Simeone, and A. Nallanathan, "Cooperative deep reinforcement learning for multiple-group NB-IoT networks optimization," in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 8424–8428.
[33] M. Jamshidi, S. S. A. Poor, N. N. Qader, M. Esnaashari, and M. R. Meybodi, "A Lightweight Algorithm against Replica Node Attack in Mobile Wireless Sensor Networks using Learning Agents," IEIE Transactions on Smart Processing & Computing, vol. 8, no. 1, pp. 58–70, 2019.
[34] H. Leopold, H. van der Aa, J. Offenberg, and H. A. Reijers, "Using Hidden Markov Models for the accurate linguistic analysis of process model activity labels," Information Systems, vol. 83, pp. 30–39, 2019.
[35] T. Chen, S. Barbarossa, X. Wang, G. B. Giannakis, and Z.-L. Zhang, "Learning and Management for Internet of Things: Accounting for Adaptivity and Scalability," Proceedings of the IEEE, vol. 107, no. 4, pp. 778–796, 2019.
[36] L. Velasco and D. Rafique, "Fault Management Based on Machine Learning," in Optical Fiber Communication Conference, 2019, pp. W3G–3.
[37] P. Wei, "Towards the Next Generation Airline Revenue Management: A Deep Reinforcement Learning Approach to Seat Inventory Control and Overbooking," arXiv preprint arXiv:1902.06824, 2019.
[38] X. Zhao, S. Ding, Y. An, and W. Jia, "Applications of asynchronous deep reinforcement learning based on dynamic updating weights," Applied Intelligence, vol. 49, no. 2, pp. 581–591, 2019.
[39] Shafik, Wasswa, Mojtaba Matinkhah, and Mamman Nur Sanda. "Network Resource Management Drives Machine Learning: A Survey and Future Research Direction." Journal of Communications Technology, Electronics and Computer Science 30 (2020): 1-15.
[40] J. Heyn, P. Gümbel, P. Bobka, F. Dietrich, and K. Dröder, "Application of artificial neural networks in force-controlled automated assembly of complex shaped deformable components," Procedia CIRP, vol. 79, pp. 131–136, 2019.
[41] F. Aznar, M. Pujol, and R. Rizo, "Obtaining fault tolerance avoidance behavior using deep reinforcement learning," Neurocomputing, 2019, pp 77-91.
[42] T. Narendra, M. S. Athulya, and P. S. Sathidevi, "Classification of Pitch Disguise Level with Artificial Neural Networks," in 2019 International Conference on Communication and Signal Processing (ICCSP), 2019, pp. 0631–0635.
[43] F. Cauteruccio et al., "Short-long term anomaly detection in wireless sensor networks based on machine learning and multi-parameterized edit distance," Information Fusion, vol. 52, pp. 13–30, 2019.
[44] S. Madjiheurem and L. Toni, "Representation Learning on Graphs: A Reinforcement Learning Application," arXiv preprint arXiv:1901.05351, 2019.
[45] J. A. Carvajal Soto, F. Tavakolizadeh, and D. Gyulai, "An online machine learning framework for early detection of product failures in an Industry 4.0 context," International Journal of Computer Integrated Manufacturing, pp. 1–14, 2019.
[46] M. M. Aburas, M. S. S. Ahamad, and N. Q. Omar, "Spatio-temporal simulation and prediction of land- use change using conventional and machine learning models: a review," Environmental monitoring and assessment, vol. 191, no. 4, p. 205, 2019.
[47] A. Enami, J. A. Torkestani, and A. Karimi, "Resource selection in computational grids based on learning automata," Expert Systems with Applications, vol. 125, pp. 369–377, 2019.
[48] A. Muñoz, J. Toutouh, and F. Jaime, "A Review of Dynamic Verification of Security and Dependability Properties," in Artificial Intelligence and Security Challenges in Emerging Networks, IGI Global, 2019, pp. 162–187.
[49] F. Zantalis, G. Koulouras, S. Karabetsos, and D. Kandris, "A Review of Machine Learning and IoT in Smart Transportation," Future Internet, vol. 11, no. 4, p. 94, 2019.
[50] R. Kashyap, "Deep Learning: An Application in Internet of Things," in Computational Intelligence in the Internet of Things, IGI Global, 2019, pp. 130– 158.
[51] T. Chilimbi, Y. Suzue, J. Apacible, and K. Kalyanaraman, "Project adam: Building an efficient and scalable deep learning training system," in 11th ${$USENIX$}$ Symposium on Operating Systems Design and Implementation (${$OSDI$}$ 14), 2014, pp. 571–582.
[52] A. M. Saghiri, M. D. Khomami, and M. R. Meybodi, Intelligent Random Walk: An Approach Based on Learning Automata. Springer, 2019.
[53] A. Enami, J. A. Torkestani, and A. Karimi, "Resource selection in computational grids based on learning automata," Expert Systems with Applications, vol. 125, pp. 369–377, 2019.
[54] B. Bordel, R. Alcarria, and D. Sánchez-de-Rivera, "A Two-Phase Algorithm for Recognizing Human Activities in the Context of Industry 4.0 and Human- Driven Processes," in World Conference on Information Systems and Technologies, 2019, pp. 175–185.
[55] Z. Bouyahia, H. Haddad, N. Jabeur, and A. Yasar, "A two-stage road traffic congestion prediction and resource dispatching toward a self-organizing traffic control system," Personal and Ubiquitous Computing, pp. 1–12, 2019.
[56] A. Karami, "An anomaly-based intrusion detection system in presence of benign outliers with visualization capabilities," Expert Systems with Applications, vol. 108, pp. 36–60, 2018.
[57] J. Li et al., "A Traffic Prediction Enabled Double Rewarded Value Iteration Network for Route Planning," IEEE Transactions on Vehicular Technology, 2019, vol. 68, pp 4170 – 4181.
[58] S. Zhang, Z. Kang, Z. Zhang, C. Lin, C. Wang, and J. Li, "A Hybrid Model for Forecasting Traffic Flow: Using Layerwise Structure and Markov Transition Matrix," IEEE Access, vol. 7, pp. 26002–26012, 2019.
[59] A. S. Hsu, J. B. Martin, A. N. Sanborn, and T. L. Griffiths, "Identifying category representations for complex stimuli using discrete Markov chain Monte Carlo with people," Behavior research methods, pp. 1–11, 2019.
[60] Y. Zou, W. Zhang, W. Weng, and Z. Meng, "Multi- Vehicle Tracking via Real-Time Detection Probes and a Markov Decision Process Policy," Sensors, vol. 19, no. 6, p. 1309, 2019.
[61] A. Berger and F. Maly, "Speech Activity Detection for Deaf People: Evaluation on the Developed Smart Solution Prototype," in Asian Conference on Intelligent Information and Database Systems, 2019, pp. 55–66.
[62] N. Jain and S. Rastogi, "Speech Recognition Systems–A Comprehensive Study Of Concepts And Mechanism," Acta Informatica Malaysia (AIM), vol. 3, no. 1, pp. 1–3, 2019.

[63] E. Lin, Q. Chen, and X. Qi, “Deep Reinforcement Learning for Imbalanced Classification,” arXiv preprint arXiv:1901.01379, 2019.
[64] A. Rao and M. Diamond, “Deep Learning of Markov Model Based Machines for Determination of Better Treatment Option Decisions for Infertile Women,” bioRxiv, p. 606921, 2019.
[65] S. S. Oyewobi, G. P. Hancke, A. M. Abu-Mahfouz, and A. J. Onumanyi, “An Effective Spectrum Handoff Based on Reinforcement Learning for Target Channel Selection in the Industrial Internet of Things,” Sensors, vol. 19, no. 6, p. 1395, 2019.
[66] E. Kayir and H. Hilal, “Q-Learning Based Failure Detection and Self-Recovery Algorithm for Multi- Robot Domains,” Elektronika ir Elektrotechnika, vol. 25, no. 1, pp. 3–7, 2019.
[67] X. Lin, R. Gu, H. Li, and Y. Ji, “A service reconfiguration scheme for network restoration based on reinforcement learning,” in 17th International Conference on Optical Communications and Networks (ICOCN2018), 2019, vol. 11048, p. 110481K.
[68] F.-C. Ghesu et al., “Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans,” IEEE transactions on pattern analysis and machine intelligence, vol. 41, no. 1, pp. 176–189, 2019.
[69] D. Terada, and C. Guo, “Automatic collision avoidance of multiple ships based on deep Q- learning,” Applied Ocean Research, vol. 86, pp. 268– 288, 2019.
[70] X. Lin, R. Gu, H. Li, and Y. Ji, “A service reconfiguration scheme for network restoration based on reinforcement learning,” in 17th International Conference on Optical Communications and Networks (ICOCN2018), 2019, vol. 11048, p. 110481.
[71] Hodo, E., Bellekens, X., Hamilton, A., Dubouilh, P. L., Iorkyase, E., Tachtatzis, C., & Atkinson, R. (2016, May). Threat analysis of IoT networks using artificial neural network intrusion detection system. In 2016 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-6). IEEE.
[72] H. H. Pajouh, R. Javidan, R. Khayami, D. Ali, and K.-K. R. Choo, “A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks,” IEEE Transactions on Emerging Topics in Computing, 2016, vol. 7, pp 314 – 323.
[73] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context-aware computing for the internet of things: A survey," IEEE communications surveys & tutorials, vol. 16, no. 1, pp. 414–454, 2014.
[74] Y. Zhu et al., “Target-driven visual navigation in indoor scenes using deep reinforcement learning,” in 2017 IEEE international conference on robotics and automation (ICRA), 2017, pp. 3357–3364.
[75] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, and J. Lloret, “Conditional variational autoencoder for prediction and feature recovery applied to intrusion detection in iot,” Sensors, vol. 17, no. 9, p. 1967, 2017.
[76] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in European conference on computer vision, 2014, pp. 818–833.
[77] Szegedy, Christian, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. "Intriguing properties of neural networks." arXiv preprint arXiv:1312.6199 (2013).
[78] J. E. Villaverde, D. Godoy, and A. Amandi, “Learning styles’ recognition in e-learning environments with feed-forward neural networks,” Journal of Computer Assisted Learning, vol. 22, no. 3, pp. 197–206, 2006.
[79] H. H. Pajouh, R. Javidan, R. Khayami, D. Ali, and K.-K. R. Choo, “A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks,” IEEE Transactions on Emerging Topics in Computing, 2016, vol.7, pp 314 – 323.
[80] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: a new learning scheme of feedforward neural networks,” Neural networks, vol. 2, pp. 985–990, 2004.
[81] N. B. Karayiannis and G. W. Mi, “Growing radial basis neural networks: Merging supervised and unsupervised learning with network growth techniques,” IEEE Transactions on Neural networks, vol. 8, no. 6, pp. 1492–1506, 1997.
[82] P. V. Krishna, S. Misra, D. Joshi, and M. S. Obaidat, “Learning automata based sentiment analysis for recommender system on cloud,” in 2013 International Conference on Computer, Information and Telecommunication Systems (CITS), 2013, pp. 1– 5.
[83] D. Hakkani-Tür, G. Riccardi, and A. Gorin, “Active learning for automatic speech recognition,” in 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002, vol. 4, pp. IV–3904.
[84] S. Misra, P. V. Krishna, V. Saritha, H. Agarwal, and Ahuja, “Learning automata-based multi- constrained fault-tolerance approach for effective energy management in smart grid communication network,” Journal of Network and Computer Applications, vol. 44, pp. 212–219, 2014.
[85] S.-H. Zahiri, “Learning automata-based classifier,” Pattern Recognition Letters, vol. 29, no. 1, pp. 40– 48, 2008.
[86] B. Braune, S. Diehl, A. Kerren, and R. Wilhelm, “Animation of the generation and computation of finite automata for learning software,” in International Workshop on Implementing Automata, 1999, pp. 39–47.
[87] K. S. Narendra and M. A. Thathachar, “Learning automata-a survey,” IEEE Transactions on systems, man, and cybernetics, no. 4, pp. 323–334, 1974.
[88] A. K. Ghosh, C. Michael, and M. Schatz, “A real- time intrusion detection system based on learning program behavior,” in International Workshop on Recent Advances in Intrusion Detection, 2000, pp. 93–109.
[89] C. L. Giles, C. B. Miller, D. Chen, H.-H. Chen, G.-Z. Sun, and Y.-C. Lee, “Learning and extracting finite state automata with second-order recurrent neural networks,” Neural Computation, vol. 4, no. 3, pp. 393–405, 1992.
[90] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S.C. de Alvarenga, “A survey of intrusion detection in Internet of Things,” Journal of Network and Computer Applications, vol. 84, pp. 25–37, 2017.
[91] A. Abduvaliyev, A.-S. K. Pathan, J. Zhou, R. Roman, and W.-C. Wong, “On the vital areas of intrusion detection systems in wireless sensor networks,” IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1223–1237, 2013.
[92] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, “Deep learning and its applications to machine health monitoring: A survey,” arXiv preprint arXiv:1612.07640, 2016.
[93] S. Misra, P. V. Krishna, H. Agarwal, A. Saxena, and M. S. Obaidat, “A learning automata-based solution for preventing distributed denial of service in Internet of things,” in 2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing, 2011, pp. 114–122.
[94] M. Weisman et al., “Machine Learning and Data Mining for IPv6 Network Defence,” in International Conference on Cyber Warfare and Security, 2018, pp. 681–XVI.
[95] W. Jiang, C.-L. Zhao, S.-H. Li, and L. Chen, “A new learning automata-based approach for online tracking of event patterns,” Neurocomputing, vol. 137, pp. 205–211, 2014.
[96] S. Raza, L. Wallgren, and T. Voigt, “SVELTE: Real- time intrusion detection in the Internet of Things,” Ad hoc networks, vol. 11, no. 8, pp. 2661–2674, 2013.
[97] Shafik, Wasswa, S. Mojtaba Matinkhah, and Mohammad Ghasemzadeh. "A Fast Machine Learning for 5G Beam Selection for Unmanned Aerial Vehicle Applications." Information Systems & Telecommunication: 262, 2019.
[98] Z. Yan, P. Zhang, and A. V. Vasilakos, “A survey on trust management for Internet of Things,” Journal of network and computer applications, vol. 42, pp. 120– 134, 2014.
[99] M. A. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, and M. Guizani, “A survey of machine and deep learning methods for internet of things (IoT) security,” arXiv preprint arXiv:1807.11023, 2018.
[100] K. Zaheer, M. Othman, M. H. Rehmani, and T. Perumal, “A Survey of Decision-Theoretic Models for Cognitive Internet of Things (CIoT),” IEEE Access, vol. 6, pp. 22489–22512, 2018.
[101] L. Cao, G. Weiss, and S. Y. Philip, “A brief introduction to agent mining,” Autonomous Agents and Multi-Agent Systems, vol. 25, no. 3, pp. 419–424, 2012.
[102] O. B. Sezer, E. Dogdu, and A. M. Ozbayoglu, “Context-aware computing, learning, and big data in internet of things: a survey,” IEEE Internet of Things Journal, vol. 5, no. 1, pp. 1–27, 2018.
[103] C. Gomez, A. Shami, and X. Wang, “Machine Learning Aided Scheme for Load Balancing in Dense IoT Networks,” Sensors, vol. 18, no. 11, p. 3779, 2018.
[104] F. M. Al-Turjman, “Information-centric sensor networks for cognitive IoT: an overview,” Annals of Telecommunications, vol. 72, no. 1–2, pp. 3–18, 2017.

[105] M. A. Alsheikh, S. Lin, D. Niyato, and H.-P. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," IEEE Communications Surveys & Tutorials, vol. 16, no. 4, pp. 1996–2018, 2014.
[106] K. Ye, "Key Feature Recognition Algorithm of Network Intrusion Signal Based on Neural Network and Support Vector Machine," Symmetry, vol. 11, no. 3, p. 380, 2019.
[107] J. Abreu, L. Fred, D. Macêdo, and C. Zanchettin, A. Rezvanian, B. Moradabadi, M. Ghavipour, M. M. D. Khomami, and M. R. Meybodi, "Wavefront Cellular Learning Automata: A New Learning "Hierarchical Attentional Hybrid Neural Networks for Document Classification, " arXiv preprint arXiv: 1901 . 06610, 2019.
[108] K. Wang, "Network data management model based on Naïve Bayes classifier and deep neural networks modelling and data-driven techniques for systems analysis," Journal of Intelligent Information Systems, in heterogeneous wireless networks," Computers & Electrical Engineering, vol. 75, pp. 135–145, 2019.
[109] P. F. Fantoni, "A neuro-fuzzy model applied to full range signal validation of PWR nuclear power plant data," INTERNATIONAL JOURNAL OF GENERAL SYSTEM, vol. 29, no. 2, pp. 305–320, 2000.
[110] M. Kahng, N. Thorat, D. H. P. Chau, F. B. Viégas, and M. Wattenberg, "GAN Lab: Understanding Complex Deep Generative Models using Interactive Visual Experimentation," IEEE transactions on visualization and computer graphics, vol. 25, no. 1, pp. 310–320, 2019.
[111] D. Popa, F. Pop, C. Serbanescu, and A. Castiglione, "Deep learning model for home automation and energy reduction in a smart home environment platform," Neural Computing and Applications, pp. 1–21, 2019.
[112] S. Baruah, "Botnet Detection: Analysis of Various Techniques," International Journal of Computational Intelligence & IoT, vol. 2, no. 2, 2019, pp 7-14.
[113] A. Mollalo, L. Mao, P. Rashidi, and G. E. Glass, "A GIS-Based Artificial Neural Network Model for Spatial Distribution of Tuberculosis across the Continental United States," International journal of environmental research and public health, vol. 16, no. 1, p. 157, 2019.
[114] R. V. McCarthy, M. M. McCarthy, W. Ceccucci, and L. Halawi, "Predictive Models Using Neural Networks," in Applying Predictive Analytics, Springer, 2019, pp. 145–173.
[115] A. Rezvanian, B. Moradabadi, M. Ghavipour, M. M. D. Khomami, and M. R. Meybodi, "Introduction to Learning Automata Models," in Learning Automata Approach for Social Networks, Springer, 2019, pp. 1– 49.
[116] W. Shafik, S. M. Matinkhah, and M. Ghasemzadeh, "Internet of Things-Based Energy Management, Challenges, and Solutions in Smart Cities," J. Commun. Technol. Electron. Comput. Sci., vol. 27, pp. 1–11, 2020.
[117] W. Shafik and S. A. Mostafavi, "Knowledge Engineering on Internet of Things through Reinforcement Learning," Int. J. Comput. Appl., vol. 975, p. 8887.
[118] S. M. Matinkhah, W. Shafik, and M. Ghasemzadeh, "Emerging Artificial Intelligence Application: Reinforcement Learning Issues on Current Internet of Things," in 2019 16th international Conference in information knowledge and Technology (ikt2019), p. 2019.
[119] S. Mostafavi and W. Shafik, "Fog Computing Architectures, Privacy and Security Solutions," J. Commun. Technol. Electron. Comput. Sci., vol. 24, pp. 1–14, 2019.
[120] W. Shafik, M. Matinkhah, M. Asadi, Z. Ahmadi, and Z. Hadiyan, "A Study on Internet of Things Performance Evaluation," J. Commun. Technol. Electron. Comput. Sci., vol. 28, pp. 1–19, 2020.
[121] S. Mostafavi and W. Shafik, "Fog Computing Architectures, Privacy and Security Solutions," J. Commun. Technol. Electron. Comput. Sci., vol. 24, pp. 1–14, 2019..