Applications of Artificial Intelligence and Machine Learning in the Area of SDN and NFV: A Survey

Anteneh A. Gebremariam*, Muhammad Usman[†], Marwa Qaraqe[†]
*Department of Information Engineering & Computer Science (DISI), University of Trento, Trento, Italy

†Information and Computing Technology, College of Science and Engineering,

Hamad Bin Khalifa University (HBKU), Education City, 34110 Doha, Qatar

Email: anteneh.gebremariam@unitn.it, {musman, mqaraqe}@hbku.edu.qa

Abstract—Artificial Intelligence (AI) and Machine Learning (ML) have gained a huge interest from academia and industry in solving very complex problems in several fields. In this paper, we present a short survey of the main application areas of AI/ML in SDN and NFV based networks. We classify the main advancements in the area in different categories based on their application track and identify the corresponding AI techniques utilized. In addition, identify and discuss the main challenges and future directions in the area. We stress that AI/ML can play a vital role in providing a way towards self-configured, self-adaptive and self-managed networks. However, the research is limited due the identified challenges in this area.

Index Terms—5G networks, Machine learning, Deep learning, SDN, NFV, Data analytics, Network planning, Network security, Network management and operations

I. INTRODUCTION

Software Defined Networking (SDN) and Network Functions Virtualization (NFV) are emerging network technologies, which are designed to decouple network services from the hardware they are executed upon. Such technologies are designed to accommodate dynamic Quality of Service (QoS) requirements, such as diversity in latency, throughput and bandwidth, onto a single physical network. This is achieved by network slicing, a technique to achieve extreme flexibility and on demand service oriented network deployments wherein a single slice is optimized to serve a single application with specific QoS requirements [1]. This results in creation of a large number software modules across the network, which need to be managed appropriately to ensure QoS. This further creates a problem of network management, since technically, it is not possible for a human operator to manage a large number of interconnected network components.

In addition, the centralized nature of SDN/NFV networks generate serious security concerns originating from single point of failure. In case the central controller is compromised, the adversary can eavesdrop or modify all the communication. Consequently, SDN/NFV networks needs special attention to ensure the availability, confidentiality, integrity and privacy of all the network nodes including the central controller.

Some other potential challenges in the virtualized networks are: network planning, resource utilization, load balancing, fault detection and management. Artificial Intelligence (AI) and Machine Learning (ML) are regarded as useful tools and potential solutions to make networks self-aware, self-adaptive,

self-secured and self-managed by incorporate intelligence in the networks. Particularly, AI and ML allow networks to learn from experience to make them more robust against vulnerabilities and failures and improve performance by making network management and operation independent of human involvement.

As a panacea, AI and ML are extensively studied in the literature to solve the aforementioned problems in general purpose communication networks. For a detailed overview of the potential solutions, one may refer to the works presented in [2]–[5]. However, in the context of SDN and NFV, a few works exist in literature that focus on the applications of AI/ML in SDN and NFV in the framework of solving the aforementioned problems. This paper presents a comprehensive survey of the existed literature in this field. This is a short survey because of limited number of the works in the area of SDN/NFV and AI/ML. In addition, in this paper, we discuss potential challenges faced by AI/ML in this area.

To the best of our knowledge, this is the first survey of the applications of AI/ML in the framework of SDN and NFV. The main contributions of this paper include:

- (i) We provide a comprehensive survey of the literature of the AI/ML to address potential challenges faced by various aspects of SDN and NFV.
- (ii) We classify different ways of applying AI/ML towards SDN and NFV architectures. Particularly, we focus on network architecture, network management, network security, and network planning.
- (iii) In each application track, we identified the machine learning technique used in the literature. Table I presents a summary of the proposed classification.
- (iv) We provide a discussion on the potential challenges and future directions in this newly emerging area.

The remainder of this paper is organized as follows. Section II presents the classification of relevant works in the area. In Section III the challenges and future directions are explained. Section IV concludes the paper.

II. APPLICATION OF AI IN SDN & NFV NETWORKS

The applications of AI and ML in SDN/NFV-based networking can be mainly classified into five broad areas.

• Network Architecture

- Network Planning / Load Balancing and Resource Utilization
- Fault/Failure Detection and Management
- Network Management and Operations
- Network Security and Breach Detection

We divide all the existed literature in the aforementioned areas and discuss them under the respective category. Table I presents a summary of the literature. In what follows, we discuss each category one by one.

A. Network Architecture

In [6] the authors propose one of the contending AI-based communication network frameworks, called Future Intelligent NEtwork-FINE, which they illustrate via SDN/NFV collaborative network example. The framework is divided into three different planes, i) intelligence plane, ii) agent plane and iii) business planes, which interact via different interfaces. The intelligent plane is the brain of the network, which is based on AI. Furthermore, it is shown that the FINE framework is feasible to be used in real communication networks and services.

B. Network Planning / Load Balancing and Resource Utilization

Tomoyuki O. *et al.* [7], propose the possibilities of applying AI technology to network operations by describing some use cases. The main use cases were targeted in tackling of issues of mobile network operations like: i) optimal network resource planning – traffic planning, ii) network resource reallocation and iii) network maintenance – network abnormality detection.

In [8], [9], the authors propose to use machine learning for assisting network operators in planning dense heterogeneous networks. Particularly, they use the data generated from current 4G networks, such as radio measurements, to develop a correlative statistical model for QoS-based network planning. The authors combine multiple learners and build ensemble methods to use them for regression in order to predict QoS in the network planning. Although, the authors focus on general 5G networks for dense heterogeneous environments but discuss SDN/NFV as an enabler of future 5G networks.

C. Fault/Failure Detection and Management

The authors in [10] presente an AI-driven malfunction detection in NFV applications via a semi-supervised learning algorithm. They did set-up a test-bed to create the training dataset for evaluating their system. The goodness of the proposed algorithm is also evaluated using the square of the standard Euclidean distance.

In [11], the first proof-of-concept on AI-assisted automated network operation system, based on failure prediction by AI, with an internetworked SDN/NFV orchestrators and operation support systems has been showcased in Mobile World Congress 2016 in Barcelona.

Sihem et al. [12] propose a global fault management framework (based on an open source solutions) called LUMEN to

provide availability and reliability of virtualized 5G end-toend service chain. LUMEN is a four-step architecture thus it is divided into four planes: i) source plane - collects data from all entities, ii) Sink plane - unifies the collected data into a single format, iii) Extraction plane - extracts the data from the sink plane and iv) Decision plane - to have an educated guess/prediction using Bayesian networks and machine learning techniques.

In [13], the authors propose NFV/SDN based self-healing and self-organizing network by incorporating intelligence in the network using machine-learning techniques. The system is a part of the European H2020 SELFNET project wherein a huge number of sensors are deployed across the network to monitor and report potential faults. The system takes intelligent decisions by using sensors and actuators to detect the problem and solve it. More specifically, their network management framework has the capabilities of self-healing, self-protection and self-optimization.

D. Network Management and Operations

In [14] a new network operation and control paradigm is proposed. This paradigm, called *Knowledge-Defined Networking (KDN)*, is based on SDN, Network Analytics (NA) and AI. The KDN is divided into four different functional planes: data plane, control plane, Knowledge Plane (KP) and management plane. The KP applies machine-learning and deep learning techniques to transform the network analytics that is collected by the management plane into knowledge, and uses this knowledge to make decisions. To be specific, three different approaches are considered: i) supervised learning, ii) unsupervised learning and iii) reinforcement learning.

Imen et al. [15] propose cognitive management architecture for managing 5G networks. They consider two use cases based on machine learning techniques: i) Service-Level Agreement (SLA) enforcement - to tackle the SLA management and ii) Mobile Quality Predictor (MQP) - to accurately predict real-time bandwidth requirement of each mobile subscriber. The cognitive architecture is made up of two main parts, the NFV framework and the cognitive framework architecture. The cognitive architecture relies on monitoring the agents that are distributed within the NFV elements. It consists of a number of blocks: data collector, data storage, Cognitive Smart Engine (CSE) and policy manager. The CSE is the core component wherein the whole intelligence resides; a special type of Artificial Neural Network (ANN) called Recurrent Neural Network (RNN) is used for implementing the SLA enforcement use-case and a supervised learning technique called Random Forest algorithm is used for implementing the MQP use-case.

Zhao *et al.* [16] present a survey of autonomic communication in the context of SDN/NFV and discuss the challenges in the network security, operations and business support incurred by the concept of autonomous and self-organized networks. Particularly, the paper focuses on automatic testing, integration and network functions deployment. In addition, the authors discuss how autonomic communication adapts to real time

TABLE I: AI Deployment with respect to applications in SDN and NFV

Reference	Application Track	AI/ML Technique Used
Xu et al. [6]	Network Architecture	A wide range of algorithms (Deep learning-ANN)
Tomoyuki et al. [7]		Anomaly detection - Unsupervised learning - Classification
Moysen et al. [8]	Network Planning/ Load Balancing	Regression - A wide range of algorithms
Moysen et al. [9]		Regression - A wide range of algorithms
KDDI R&D Lab [11]	- Fault/Failure Detection and Management	Not mentioned but it seems like a classification problem
Sihem et al. [12]		Classification - Bayesian networks
Julian et al. [10]		Classification/Clustering - Autoencoders
Perez et al. [13]		Not mentioned but it seems like a classification problem
Albert et al. [14]	Network Management and Operations	A wide range of algorithms from supervised, unsupervised and reinforcement learning
Imen <i>et al.</i> [15]		ANN (RNN) and random forest
Zhao <i>et al</i> . [16]		A wide range of algorithms from supervised, unsupervised and reinforcement learning
Moysen et al. [17]		A wide range of algorithms from supervised, unsupervised and reinforcement learning
Mestres [18]		Regression - ANN
Vergara et al. [19]		Classification - Naive Bayes (NB)
He <i>et al.</i> [20]		Classification - A wide range of algorithms (K nearest neighbors, SVM, etc.)
Zorzi et al. [21]		GDNN
Jaafar et al. [22]		ANN
Ankur et al. [23]		A dynamic game with multiple players, based on Nash Folk Theorem
Assis et al. [24]	Network Security and Breach Detection	Game theory
Richard et al. [25]		Game theory and NB classifier
Pan et al. [26]		A wide range of algorithms from deep learning
Gardikis et al. [27]		A wide range of algorithms from supervised, unsupervised and reinforcement learning

threats for providing self-managed protection against them, utilizing SDN/NFV security defenses.

Another work that investigates the use of machine learning in SDN/NFV for network organization and management is presented in [17]. The authors mainly focus on the autonomous operations, administration and management of the network in future 5G networks. However, the authors partially discuss the SDN/NFG as an enabler of 5G networks for integrating various applications of diverse QoS requirements and the use of machine learning to implement autonomous adaptability and take an advantage of the experience during decision process. The authors elaborate supervised learning for estimation, prediction and classification of different variables for network management, unsupervised learning to identify anomalous behavior and reducing dimensions of the network data and reinforcement learning for the cases when network management require network parameter control.

The author in [18] proposes incorporating ANN to Autoregressive Integrated Moving Average (ARIMA) models for traffic forecasting and network modeling. The author provides different use-cases and applications to exploit SDN, network analysis (NA) and ML to describe new paradigms of future networks. Focusing on NFV, the author models the performance of single Virtual Network Function (VNF) as a function of input traffic. Particularly, the author estimates the CPU consumption of a VNF using input traffic characteristics.

The work in [19] proposes ML algorithms to manage the hidden traffic in NFV-based networks. Hidden traffic in virtualized networks is the invisible traffic that does not hit the physical layer. The authors benchmark the performance of different supervised learning algorithms for the classification of IP traffic in NFV-based networks. Based on the demonstrated results, the authors find Naive Bayes (NB) algorithm as the best IP traffic classifier in NFV-based networks, with a

maximum value of 99.9%. Similar kind of work is presented in [20].

In [21], the authors introduce a system concept of COgnition-BAsed NETworkS (COBANETS), with SDN tools as building blocks, for the design and operation of next generation communication networks. In particular, the authors propose to use unsupervised deep learning algorithms with network virtualization paradigms to incorporate the automatic optimization and reconfiguration strategies at the system level.

E. Network Security and Breach Detection

The authors in [22] use ANN to design and evaluate a framework for cognitive SLA enforcement of networking services involving VNFs and SDN controllers. The framework identifies and detects a possible Service Level Objective (SLO) breaches such that facilitating a proper management actions for service provides. They evaluate their framework via a testbed implementation applied to a streaming service running on top of a virtualized and softwarized infrastructure (NFV and SDN). The ANN solves a classification problem in order to determine the presence of either SLO breach or No SLO breach. They mainly focused two types of ANNs, Feed-Forward Neural Network (FFNN) and the Recurrent Neural Network (RNN).

The authors in [24] propose a proactive defense system against Denial of Service (DoS) and Distributed Denial of Service (DDoS) attack in SDNs. The system is based on game theoretic approach using holt-winters and genetic algorithm with fuzzy logic. Their proposed system, Holt - Winters for Digital Signature (HWDS), unites the problems of anomaly detection and identification with a game theory based decision model. Ankur *et al.* [23] proposed a game theoretic attack analysis (DDoS attacks) and countermeasure selection model in SDN based network environments. The model is based on reward and punishment in a dynamic game with multiple

players where the network bandwidth of the attackers is downgraded for a certain period of time and then restored to normal when the player resumes to cooperate. A Nash Folk Theorem is used to implement the reward and punishment of the players based on their respective actions, cooperation and attackers who are part of DDoS attack. In [25], a game theory and Machine Learning (ML - NB classifier) are combined to model attacker's behavior in ML feature space. This work uses spam filtering as a target application to provide defense against current and future attacks.

The authors in [26], discuss cybersecurity challenges and potential opportunities in edge-computing environments for IoT in the framework of SDN and NFV. The authors discuss how ML and AI can potentially help to resolve these challenges. Particularly, the authors propose to use Virtual Machines (VMs) as firewalls against security attacks and/or intrusion detection. Additionally, the authors discuss how ML and AI can be utilized to analyze cyber behaviors and identify potential threats and vulnerabilities in SDN/NFV based edge-computing environments for IoT. The authors propose to deploy automatic robots that can potentially scan an organization environment adopting deep learning technology to identify any suspicious behavior within the network.

The European Project SHIELD [27] focuses to protect networks against intruders and other threats using NFV-enabled environments. The project elaborates the prevention and detection of various security vulnerabilities using machine learning algorithms trained by the security experts. In addition, the authors survey the available cybersecurity technologies focusing on virtualized services on the cloud, such as AlienVault, ArcadiaData, BlackStratus and others.

III. DISCUSSION

In this section we point out the main challenges and future directions in applying AI/ML in SDN and NFV network environments. We divide the main challenges into four main categories: i) computational complexity and latency, ii) resource requirement iii) access to resources - datasets and iv) storage of valuable information.

A. Computational Complexity and Latency

Most of the cloud-based operations that require ML involve quite a lot number of computationally complex operations and algorithms, thus leading to higher latency in performing the required operations. For time sensitive operations, there should be ways to perform some of the operations with more efficient algorithms or even implementing some of the modules close to the edge of the network. Future works should focus in analyzing and quantifying the impacts of complexity and latency in the system performance.

B. Resource Requirement (CPU/Memory)

As a consequence of the computational complexity described in Section III-A, the allocated resources at the cloud need to be capable of performing the required operations. A fast CPU resources and memory requirement should be

allocated in order to compute complex operations with low latency and store a huge amount of datasets, respectively. Even though, we assume the centralized data centers could provide enough resources, it is always a good practice to allocate the required resources according to the need. This in turn saves unnecessary power wastage in the data centers.

C. Acess to Resources - Datasets

The other main challenge in applying AI towards telecommunication environment is the lack of openly available standard datasets. These datasets are mostly owned by mobile operators who have a strict rule of not to sharing any of their data for public use. In order to tackle this issue, we propose two possible ways of generating the required datasets for evaluating different algorithms: i) via the implementation of testbeds (e.g., [11]) and ii) via standard network simulators/emulators (e.g., [28]–[30]). Furthermore, making these datasets available for public use accelerates innovation and provides a good incentive to other researchers to share their datasets. Thus helping us to create a standardized dataset that could be used by the community to evaluate different use-cases, applications and scenarios.

D. Storage of Valuable Information

Another problem is the storage of data itself. Within communication networks, there is no trend of storing communication data for future use and experience. Only a small of data is generally stored and a lot of worthy data goes unstored after usage. There is an immense need to store the valuable data generated from different parts of the network, especially, in the case network failure and cyber attacks. This will not only benefit network operators to learn from their bad and good experiences but also benefit researchers for the experimentation of self-organized, self-managed and self-secured networks.

IV. CONCLUSIONS

This paper presents a comprehensive survey of the applications of AI/ML in the area of SDN and NFV. Particularly, we identified different application tracks of the literature and divided the literature based on those tracks. In addition, in each of the presented work, we identified the corresponding AI/ML technique utilized. Subsequently, we discussed potential challenges to conduct the AI/ML based research in different aspects of SDN and NFV. Specifically, we stress that lack of available datasets and limited access to available network resources subsequently limit the research in this emerging area.

In future, we plan to quantify the usage of computational resources of network elements by executing different ML algorithms for various different network scales and traffic loads. Specifically, we aim to find a relation between traffic load on the network and the computational complexity, depending on the scale of the network, *i.e.*, the number of network nodes involved.

REFERENCES

- [1] A. A. Gebremariam, M. Usman, P. Du, A. Nakao, and F. Granelli, "Towards E2E Slicing in 5G: A Spectrum Slicing Testbed and Its Extension to the Packet Core," in 2017 IEEE Globecom Workshops (GC Wkshps), Dec 2017, pp. 1–6.
- [2] T. T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *IEEE Communications* Surveys & Tutorials, vol. 10, no. 4, pp. 56–76, 2008.
- [3] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commu*nications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, 2016.
- [4] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Security and Privacy (SP)*, 2010 IEEE Symposium on. IEEE, 2010, pp. 305–316.
- [5] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5g: how to empower son with big data for enabling 5g," *IEEE network*, vol. 28, no. 6, pp. 27–33, 2014.
- [6] X. Guibao, M. Yubo, L. Jialiang, "INCLUSION OF ARTIFICIAL INTELLIGENCE IN COMMUNICATION NETWORKS AND SER-VICES," ITU Journal: ICT Discoveries, Special Issue, no. 1, pp. 1–6, Oct. 13 2017.
- [7] T. Otani, H. Toube, T. Kimura, and M. Furutani, "APPLICATION OF AI TO MOBILE NETWORK OPERATION," ITU Journal: ICT Discoveries, Special Issue, no. 1, pp. 1–7, Oct. 13 2017.
- [8] J. Moysen, L. Giupponi, and J. Mangues-Bafalluy, "On the potential of ensemble regression techniques for future mobile network planning," in *Computers and Communication (ISCC)*, 2016 IEEE Symposium on. IEEE, 2016, pp. 477–483.
- [9] M. Jessica and G. Lorenza and M.-B. Josep, "A machine learning enabled network planning tool," in *Personal, Indoor, and Mobile Ra*dio Communications (PIMRC), 2016 IEEE 27th Annual International Symposium on. IEEE, 2016, pp. 1–7.
- [10] Julian A., Mathias S., Lia A. and Hans D. S., "An AI-driven Malfunction Detection Concept for NFV Instances in 5G," CoRR, vol. abs/1804.05796, pp. 1–5, 2018.
- [11] KDDI Research and Development Laboratories, Inc. (2016, Feb. 22) World's first successful AI-assisted automated network operation system PoC towards 5G. [Online]. Available: http://www.kddi-research.jp/english/newsrelease/2016/022201.html
- [12] S. Cherrared, S. Imadali, E. Fabre, and G. Gössler, "LUMEN: A global fault management framework for network virtualization environments," in 2018 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN). IEEE, 2018.
- [13] M. G. Pérez, G. M. Pérez, P. G. Giardina, G. Bernini, P. Neves, J. M. Alcaraz-Calero, Q. Wang, and K. Koutsopoulos, "Self-Organizing Capabilities in 5G Networks: NFV & SDN Coordination in a Complex Use Case," 2018.
- [14] A. Mestres, A. Rodriguez-Natal, J. Carner, P. Barlet-Ros, E. Alarcón, M. Solé, V. Muntés-Mulero, D. Meyer, S. Barkai, M. J. Hibbett et al., "Knowledge-defined networking," ACM SIGCOMM Computer Communication Review, vol. 47, no. 3, pp. 2–10, 2017.
- [15] I. G. B. Yahia, J. Bendriss, A. Samba, and P. Dooze, "CogNitive 5G networks: Comprehensive operator use cases with machine learning for management operations," in *Innovations in Clouds, Internet and Networks (ICIN)*, 2017 20th Conference on. IEEE, 2017, pp. 252–250
- [16] Z. Zhao, E. Schiller, E. Kalogeiton, T. Braun, B. Stiller, M. T. Garip, J. Joy, M. Gerla, N. Akhtar, and I. Matta, "Autonomic Communications in Software-Driven Networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2431–2445, 2017.
- [17] J. Moysen and L. Giupponi, "From 4G to 5G: Self-organized Network Management meets Machine Learning," Computer Communications, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140366418300380
- [18] A. Mestres Sugrañes, "Knowledge-defined networking: a machine learning based approach for network and traffic modeling," 2017.
- [19] J. Vergara-Reyes, M. C. Martinez-Ordonez, A. Ordonez, and O. M. C. Rendon, "IP traffic classification in NFV: A benchmarking of supervised Machine Learning algorithms," in *Communications and Computing (COLCOM)*, 2017 IEEE Colombian Conference on. IEEE, 2017, pp. 1–6.

- [20] L. He, C. Xu, and Y. Luo, "VTC: Machine learning based traffic classification as a virtual network function," in *Proceedings of the 2016* ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization. ACM, 2016, pp. 53–56.
- [21] M. Zorzi, A. Zanella, A. Testolin, M. D. F. De Grazia, and M. Zorzi, "Cognition-based networks: A new perspective on network optimization using learning and distributed intelligence," *IEEE Access*, vol. 3, pp. 1512–1530, 2015.
- [22] J. Bendriss, I. G. B. Yahia, P. Chemouil, and D. Zeghlache, "AI for SLA management in programmable networks," in *DRCN 2017-Design* of Reliable Communication Networks; 13th International Conference; Proceedings of. VDE, 2017, pp. 1–8.
- [23] A. Chowdhary, S. Pisharody, A. Alshamrani, and D. Huang, "Dynamic game based security framework in sdn-enabled cloud networking environments," in *Proceedings of the ACM International Workshop on Secu*rity in Software Defined Networks & Network Function Virtualization. ACM, 2017, pp. 53–58.
- [24] M. V. De Assis, A. H. Hamamoto, T. Abrao, and M. L. Proença, "A game theoretical based system using holt-winters and genetic algorithm with fuzzy logic for DoS/DDoS mitigation on SDN networks," *IEEE Access*, vol. 5, pp. 9485–9496, 2017.
- [25] R. Colbaugh and K. Glass, "Predictability oriented defense against adaptive adversaries." In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, Korea, pp. 2721–2727, Oct. 14-17 2012.
- [26] J. Pan and Z. Yang, "Cybersecurity Challenges and Opportunities in the New Edge Computing+ IoT World," in *Proceedings of the 2018 ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization*. ACM, 2018, pp. 29–32.
- [27] G. Gardikis, K. Tzoulas, K. Tripolitis, A. Bartzas, S. Costicoglou, A. Lioy, B. Gaston, C. Fernandez, C. Davila, A. Litke et al., "SHIELD: A novel NFV-based cybersecurity framework," in Network Softwarization (NetSoft), 2017 IEEE Conference on. IEEE, 2017, pp. 1–6.
- [28] (2018, Aug. 26). [Online]. Available: https://www.nsnam.org/
- [29] (2018, Aug. 26). [Online]. Available: http://www.openairinterface.org/
- [30] (2018, Aug. 26). [Online]. Available: http://mininet.org/