Two Decades of AI4NETS - AI/ML for Data Networks: Challenges & Research Directions

Pedro Casas
AIT Austrian Institute of Technology, Vienna, Austria
pedro.casas@ait.ac.at

***Abstract*—The popularity of Artificial Intelligence (AI) – and of Machine Learning (ML) as an approach to AI, has dramatically increased in the last few years, due to its outstanding performance in various domains, notably in image, audio, and natural language processing. In these domains, AI success-stories are boosting the applied field. When it comes to AI/ML for data communication Networks (AI4NETS), and despite the many attempts to turn networks into learning agents, the successful application of AI/ML in networking is limited. There is a strong resistance against AI/ML-based solutions, and a striking gap between the extensive academic research and the actual deployments of such AI/ML-based systems in operational environments. The truth is, there are still many unsolved complex challenges associated to the analysis of networking data through AI/ML, which hinders its acceptability and adoption in the practice. In this positioning paper I elaborate on the most important show-stoppers in AI4NETS, and present a research agenda to tackle some of these challenges, enabling a natural adoption of AI/ML for networking. In particular, I focus the future research in AI4NETS around three major pillars: (i) to make AI/ML immediately applicable in networking problems through the concepts of *effective learning*, turning it into a useful and reliable way to deal with complex data-driven networking problems; (ii) to boost the adoption of AI/ML at the large scale by learning from the Internet-paradigm itself, conceiving novel distributed and hierarchical learning approaches mimicking the distributed topological principles and operation of the Internet itself; and (iii) to exploit the softwarization and distribution of networks to conceive *AI/ML-defined Networks (AIDN)*, relying on the distributed generation and re-usage of knowledge through novel *Knowledge Delivery Networks (KDNs)*.**

***Index Terms*—Machine Learning; Artificial Intelligence; Data Communication Networks; Data-driven networking; Knowledge Delivery Networks (KDNs); AI/ML-defined networking (AIDN).**

I. Introduction

The impressive success of AI/ML in multiple data-driven problems over the past decade has motivated a revamped interest in AI/ML-based solutions to networking problems. Data communication networks generate a wealth of big data, which offers the chance to characterize and analyze complex networking systems in a purely data-driven manner, from an end-to-end perspective, mapping input data to output targets, and without the need to model and characterize the functioning of specific components. With such wealth of big data, learning-based networking would represent a cornerstone to better, more efficient, more reliable, more flexible, and safer networks. The immediate question that poses then is: **why is networking not profiting from the outstanding success AI/ML is having in so many other data-driven domains?**.

Making AI/ML an accepted and fruitful approach to networking in operational networks is extremely challenging. All the major learning and big data challenges - the famous *4 Vs*, are present when dealing with operational networks' data: massive volumes of complex and heterogeneous data (**V**olume and **V**ariety), fast and highly dynamic streams of data (**V**elocity), lack of ground truth for learning (**V**eracity), highly unbalanced data, lack of visibility due to massive adoption of end-to-end encryption, problems for transparently interpreting AI/ML-based systems, and more. As a consequence, getting AI/ML to work in operational networks and at scale, in such a variety of dynamic data environments, with limited human intervention, and providing proper and reliable results, is still an open problem which needs to be solved.

Despite an extensive academic research in the application of AI/ML to data communication Networks (AI4NETS) over the past two decades, there is a striking gap in terms of actual deployments of AI/ML-based solutions: compared to other network management approaches, AI/ML is rarely employed in operational scenarios.

To understand the reasons for this limited success of AI/ML in operational networks, I first present in Section II a brief overview on the broad domain of AI/ML, considering conventional ML and more recent developments over the past decade. I then describe past and more recent applications of AI/ML to networking problems - taking ML as the most representative approach to AI when it comes to networking. In Section III I elaborate on the the major bottlenecks hindering a wide and successful adoption of AI/ML in operational networking areas. At last, in Section IV I present a potential research agenda in AI4NETS to advance and tackle some of these bottlenecks.

II. State-of-the-Art

A. Machine Learning – Past & Present

As a branch of AI or basically as a practical approach to AI, the field of ML has been studied for more than 60 years now, and today there is a plethora of ML approaches and techniques [4], [19], [28], [43], [47], covering the three main learning paradigms, namely supervised, unsupervised, and reinforcement learning. Very popular ML algorithms include Support Vector Machines (SVMs), Decision Trees, Neural Networks, Naïve Bayes, Random Forrest, Clustering approaches such as K-means and DBSCAN, Q-learning by reinforcement, and the

978-1-7281-4973-8/20/$31.00 © 2020 IEEE

list goes on. These more conventional ML algorithms fall within the so called *shallow* learning techniques, which are basically bounded to well defined input data representations (i.e., features), and are incapable to process data in its raw form. Traditional ML has been therefore very much dependent on careful feature engineering and expert domain knowledge, to derive the most relevant set of features out of the raw data under analysis.

With the major advances in computational processing capacity, notably through the massive production and adoption of Graphics Processing Units (GPUs) and more recent Tensor Processing Units (TPUs), and the surge of data availability over the past decade, shallow learning approaches started getting overtaken by a new breed of *deep* ML approaches, based on the concepts of Deep Learning (DL) [19], [28], [45]. Different from conventional ML, DL is extremely data-driven and somehow agnostic to the specific type of data, requiring very big amounts of it to learn, and operating in a completely black-box manner. The main advantage of DL is its inherent capacity to learn directly from raw input data, without requiring heavily hand-crafted features to provide good results. The key behind DL is the so-called representation learning paradigm [3], which offers a set of methods allowing a ML algorithm to automatically discover the best data representations or features from raw data inputs for the specific learning task (e.g., classification). DL methods are basically representation learning methods with multiple levels of representation or abstraction, obtained by composing simple but non-linear consecutive transformation steps or layers, each of them providing a more abstract representation of the data. DL has dramatically improved the state-of-the-art in multiple domains, including speech recognition, visual object recognition, object detection and many others such as genomics.

Worth mentioning are two milestones enabling a successful and massive application of DL to data-driven problems, led by the dubbed "Godfather of DL", Geoffrey Hinton: in 2006, Hinton et al. [22] introduced a novel and effective way to train very deep neural networks by pre-training one hidden layer at a time, using the unsupervised learning procedure for restricted Boltzmann machines [17]. In 2012, one of his students, Alex Krizhevsky, designed a deep convolutional network called the AlexNet, which strongly helped to revolutionize the field of computer vision, by almost halving the error rate for object recognition at the 2012 ImageNet challenge [27]. This precipitated the rapid adoption and popularity of DL in computer vision problems, naturally extending to other domains. Also worth mentioning are the recent developments in Deep Reinforcement Learning [18], playing a critical role in todays success of DL, notably through the popular implementations of AlphaGo, AlphaZero, and more recently AlphaStar.

But of course, even if today everything in AI/ML is about deep learning, the AI community is extremely active with other major challenges faced by the application of ML in the practice, which together would enable the next major steps in AI. Novel learning paradigms such as the major breakthrough introduced in 2014 by Ian Goodfellow – the Generative Adversarial Networks (GANs) approach, where models train to new levels of performance by *pitting* against one other, as well as new approaches to improve the applicability of AI in the practice, by alleviating the black-box effect of complex ML models and make them more easy to understand for the end user – eXplainable AI (XAI) [30], [42], all contribute to the popularity and successful application of ML. Other major challenges and disciplines include the problems of Continual or Lifelong Learning [40], Robust Learning [20], Hierarchical Learning [8], Multi-task Learning [11], Meta Learning [29], and Transfer Learning [1], among others.

B. Where are we in AI4NETS?

When it comes to the popularity of AI4NETS, the picture looks less promising. The application of AI concepts and ML approaches to networking problems has today more than two decades of existence, with the first steps taken back in the late 90's, when the concept of Cognitive networking (CN) [48] - term first coined in 1998 by researchers in KTH Sweden, was introduced. The CN paradigm describes a network with cognitive capabilities which could learn from past observations and behaviors, to better adapt to end-to-end requirements. CN has been strongly re-furbished along time, referring to it as self organizing networks, self-aware networks, self-driving networks, intelligent networks, and so forth [32]; this has motivated a surge of papers applying all kinds of ML algorithms to different networking problems at hand. Today we have a broad number of surveys [2], [5], [7], [10], [16], [24], [39] over-viewing the literature on the application of ML to diverse networking problems, including traffic prediction, traffic classification, traffic routing, congestion control, network resources management, network security, anomaly detection, QoS and QoE management, etc. A common trend we find in the existing literature is that, for the most part of the papers doing AI/ML for networking, there is a systematic lack of analysis on the multiple aspects which could lead to eventually re-use and reproduce, generalize, or even apply the obtained results in real deployments. In a nutshell, most of what we have in AI/ML applied to networking has been about grabbing one particular ML approach and testing it on some particular networking dataset - hopefully large, but in reality, of limited size and most probably of limited representability of the underlying problems.

Today we start seeing papers addressing more modern flavors of AI/ML and their application in networking, e.g., with initial results on deep learning [35], [41], [44], transfer learning [50], XAI for network security [21], deep reinforcement learning for network management [31], [33], [34], etc. However, the speed of adoption of AI/ML-based solutions to current networking problems is extremely slow, with a strong resistance from not only the applied networking domain, but also the networking research community. There is a striking gap between the extensive academic research and the actual deployments of such AI/ML-based systems in operational environments.

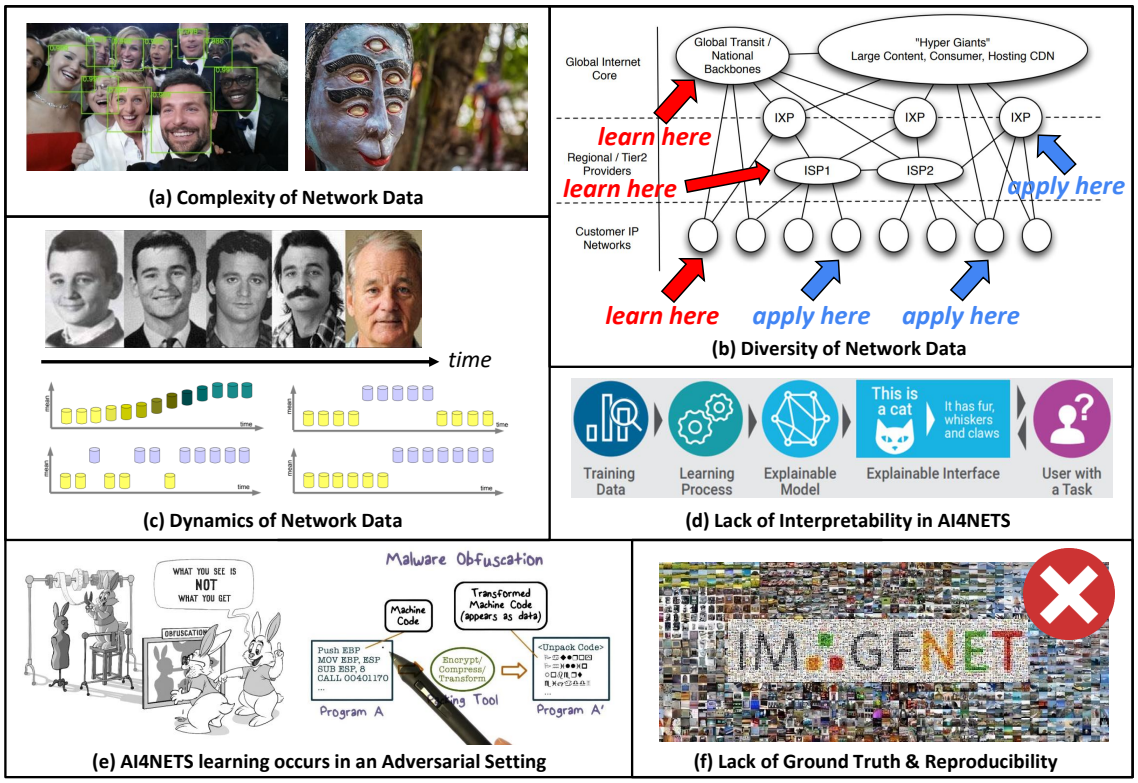

Figure 1: Why is AI/ML so challenging in networking applications?

The picture of AI/ML in networking looks even more downcast when considering the many past efforts we have seen in turning large-scale networks into cognitive agents. Some years after the introduction of CN, Clark (one of the pioneers of the Internet) et al. [12] proposed the "Knowledge Plane" (KP) for the Internet, a pervasive, distributed system within the network that would provide the required levels of abstraction and end-to-end visibility to realize the CN paradigm. While both CN and KP have inspired numerous research projects over the last two decades, today we still do not have any of these paradigms bringing solutions to real network deployments. With the more recent proliferation and developments on flexible networking, i.e., the Software Defined Networking (SDN) paradigm, and based on the growing availability of data coming from networking applications, there have been new attempts to realize the notions of CN and KP into operational deployments [37], and to make of AI an enabler for data-driven networking [23]. Besides the complexities associated to SDN and network flexibility, and the limitations on network measurement and data collection in the large scale, the truth is that there are still many unsolved complex challenges associated to the analysis of networking data through ML, which hinders its acceptability and adoption in the practice. A quick thought on this already reveals part of the problem: different from other AI/ML-boosted disciplines where data comes from "natural" sources (image, video, voice, text) or highly constrained environments (gaming), network data lives in hard to charaterize manifolds and is highly complex, both in variety and in statistical properties (highly non-linear, statistically variant, multi-structured, etc.). Not surprisingly, already a decade ago, top networking researchers showed that the Internet has characteristics of complex systems [14], making it hard to model and predict its behavior.

The first step to unleash the power of AI in networking and to turn Clark's KP vision into a reality is to tackle the fundamental problems limiting AI's successful application in the practice. This is exactly where I think the agenda in AI4NETS has to focus on: enabling a systematic and successful application and adoption of AI/ML in networking, by making AI/ML "immediately applicable" to large-scale networking problems, addressing the major bottlenecks through the application of what I have defined as **effective machine learning** and **Knowledge Delivery Networks (KDNs)**. I further elaborate on these concepts in Section IV.

III. The Challenges – Why is AI/ML Slowly Succeeding in Networking?

Let us dig deeper into the causes of such an impressive mismatch between the developments of AI/ML for networking and the popularity and success of AI in other data-driven domains. Next I provide a non-exhaustive summary on some of the main bottlenecks and challenges hindering a bright adoption of AI4NETS, including part of those previously flagged by Paxon in his seminal work on machine learning

for anomaly detection [46]. I make reference to Figure 1 as a graphical guide to the different limitations we are facing in AI4NETS.

Data Complexity (Figure 1a): as I mentioned before, the complexity (and heterogeneity) of the data related to Internet-like networks is one of the most significant bottlenecks to AI4NETS. The Internet, and in general large-scale networks, are a complex tangle of networks, technologies, applications, services, devices and end-users. As a consequence, even if networking protocols are well defined (at least a-priori), the interaction among all these components makes of the resulting data a major challenge when it comes to learning anything out of it. Quoting Eric Schmidt, former Google's CEO: "the Internet is the first thing that humanity has built that humanity does not understand, the largest experiment in anarchy that we have ever had". Data arising from domains where AI has so far shown very successful results generaly comes from more predictable and easy to undertand sources, for example natural images, or natural language, or even from well defined and controlled environments when it comes to more recent AI victories in gaming - and not only easy Atari-like games, but more complex ones such as the recently tackled Star Craft game by Google DeepMind. For example, while one would not expect to find images of faces with five eyes and three noses when addressing face-recognition tasks, networking data can change in unpredictable manners, making it harder to deal with data-driven approaches.

Diversity of Network Data (Figure 1b): besides complexity, network data often exhibits much more diversity than one would intuitively expect. Even within a single network, the network's most basic characteristics - e.g., the mix of different applications, can exhibit immense variability, rendering them unpredictable over short time intervals. With the ever-growing number of devices connected to large-scale networks and the continuous emergence of new applications and services running on them, the diversity and unpredictability of networking data can just increase.

Data Dynamics (Figure 1c): this brings us to an even more critical issue in networking; networking data is non-static, and generally comes in the form of data streams, which by nature are difficult to analyze. Networking stream data is full of constant concept drifts - changes in the underlying statistical properties, which requires different from traditional approaches to make sense and good use out of it [38].

Lack of Learning Generalization (Figure 1b): as a consequence, it becomes extremely difficult in the networking practice to learn models which can generalize to other environments, different from those where the training data comes from. A very common issue in learning for networking problems is that those solutions which are built and calibrated for certain types of networks, are no longer useful when deployed in the operational environments.

Lack of Ground Truth (Figure 1f): supervised learning needs massive amounts of labeled data to learn, but "in the wild" networking data is usually non-labeled; labeling data in operational environments is extremely costly and error prone. As a result, AI/ML-based solutions have to heavily rely on unsupervised or semi-supervised learning schemes, with their corresponding limitations in terms of knowledge generation.

Dealing with highly Imbalanced Problems: network data is usually highly imbalanced when it comes to classes representing different categories in the data; for example, network anomalies and attacks reasonably occur much less often than normal instances; this has to be properly taken into account to avoid over-fitting and other undesirable learning problems.

Lack of Standardized and Representative Datasets (Figure 1f): different from other AI/ML-related domains, where well established, publicly available datasets are available for testing, evaluation and benchmarking purposes (e.g., ImageNet in image processing), it is very difficult to find appropriate public datasets to assess AI4NETS. While one of the main reasons for this lack clearly arises from the data's sensitive nature - including end-user privacy, other limitations come from the efforts required to build proper and representative datasets in networking. Given the scale of Internet-like networks, the massive volumes of data, and the multiplicity of operational conditions, building such a representative dataset is a daunting task. So in simple words, there is no ImageNet in AI4NETS.

Lack of Model Performance Bounds & High Cost of Errors: deploying AI/ML-based solutions in operational environments comes with the associated challenge of incurring in costly errors - specially when dealing with critical applications such as security, which usually ISPs and network vendors are not willing to bear. Being data-driven by nature, and prone to outliers, it is extremely challenging to provide tight performance bounds on trained AI models. Robust learning is paramount for networking.

Learning occurs in an Adversarial Setting (Figure 1e): learning data is full of adversarial examples in many networking problems, which again makes the learning process more cumbersome; this is the case not only in network security applications, where a classic arms-race is evident, but also when it comes to traffic identification and classification tasks, as applications do not want to be tracked by intermediate entities, and therefore obfuscate and dynamically modify their functioning to bypass monitoring and avoid traffic engineering policies.

Lack of Model Transparency (Figure 1d): the lack of transparency of most ML models limits their application in real deployments, especially in critical applications such as security. If you cannot understand why this or that decision is taken, then you would not trust it and therefore not use it. The complexity of current DL models exacerbates this black-box effect, therefore limiting its application in the practice.

Lack of a combined Knowledge & Expertise in both networking and AI/ML: the networking and the AI/ML communities are totally separated, and it is very hard today to find the required expertise to properly tackle networking problems through AI/ML. So far, AI/ML in networking has been about grabbing one particular ML approach and testing it on (hopefully) large-scale datasets, without really mastering the underlying properties and requirements of the selected ML

approach. At the same time, we have seen AI researchers applying ML to networking problems which they do not really understand, and therefore the validity of the obtained conclusions and their analysis is quite limited in the best case. There is a huge need of interdisciplinary approaches and teams to tackle this limitation.

Up to date, there has not been any systematic research effort towards better understanding and tackling the main issues limiting a successful adoption of AI to data-driven networking problems. As a consequence, the lack of success-stories in AI4NETS is not surprising at all.

IV. A Research Agenda in AI4NETS

In the next lines I propose an ambitious research agenda to materialize novel ideas, concepts, and learning approaches towards **bridging the gap between AI/ML and its successful and systematic application to improve networking problems at the large scale**. The so far presented state of affairs in AI4NETS clearly suggests that resolving the main bottlenecks hindering a natural adoption of AI/ML in data-driven networking problems requires a solid grasp on the main concepts and principles behind network measurements and AI/ML. In Figure 2, I depict the three main pillars of the proposed research agenda, including (i) effective learning, (ii) Internet (distributed) learning, and (iii) AI defined Networks.

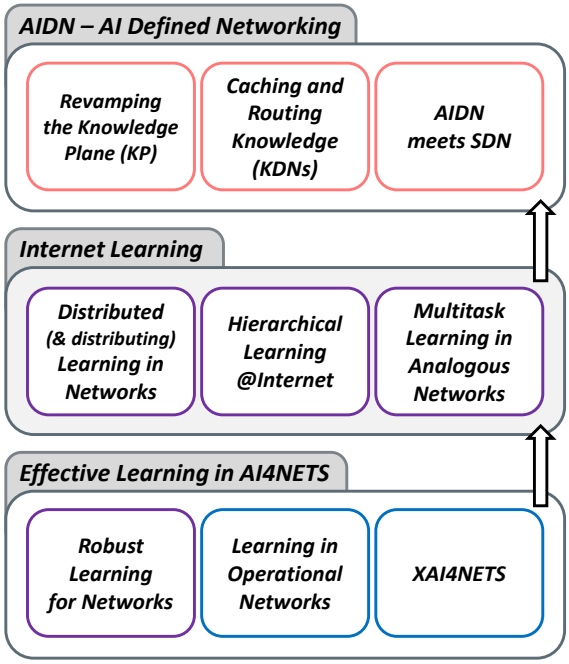

Figure 2: A research agenda in AI4NETS.

Effective Learning in AI4NETS deals with making AI/ML immediately applicable to networking problems. By immediately applicable, I mean a successful application in terms of performance, generalization of results, and model transparency, leading to a subsequent natural adoption in networking problems. The principles behind what I shall refer to as **effective learning** for networks consists of a series of general AI/ML-related problems (robust, transferable and explainable learning), applied to the particular networking challenges outlined above. The work I propose in this area can be further structured under three specific topics – **(i) Robust learning for networking:** consists of the investigation, tailoring and application of more novel learning paradigms which are currently showing great performance in other domains, including (among others) robust-deep-learning based paradigms and representation learning, (deep) reinforcement leaning for closed-loop applications, and the study of robust transfer in networking learning tasks. All these approaches have evolved driven by non-networking applications (e.g, image processing), and as such, need to be seriously re-considered when dealing with networking data. **(ii) Learning in operational networks:** should tackle those challenges which arise from the networking practice, when dealing with real networks data and when deploying AI/ML-based solutions into operational settings; learning with highly imbalanced data, learning with limited (or none) availability of labels/ground truth, or learning in highly dynamic settings and under the presence of frequent concept drifts are some of the problems tackled here. **(iii) Explainable AI for networks:** addresses one of the most important challenges stopping the adoption of AI/ML-based solutions, namely the transparency and understanding of the decisions taken by ML models in networking problems. Conceiving approaches to better explain model properties, explain the specific rationale behind decisions, and validate the operation of black-box like models, will ease their adoption in operational environments.

Internet Learning is complementary to the first pillar, but instead of asking *what can AI/ML do for the Internet*, the target here is to figure out ***what can the Internet do for AI/ML***; in particular, here I propose a radically different approach to design distributed learning algorithms, by leveraging the underlying properties of the Internet itself in terms of distributed topological principles and operation, to build better and more robust learning approaches for networking. As noted by Stonebraker - one of the gurus of big data analytics, distributed learning is the next frontier to AI/ML scalability [9]. One major bottleneck for scalable and distributed AI/ML is that most ML algorithms were not originally conceived to operate in a distributed manner, which seriously limits the availability of ML libraries for parallel computing platforms. To realize this Internet learning concept, and as noted in [6], I propose to leverage the fact that distributed network algorithms generate a wealth of big data in the form of problem instance and solution pairs as part of their normal operation, which introduces an unexplored opportunity to learn from this examples at a massive scale. The work I propose in this area can be further structured under three specific topics – **(i) Distributed (and distributing) learning in networks:** here the focus is on the principles behind distributed machine learning and distributed computation [13], [15], [49], conceiving approaches to realize such distributed learning by relying on distributed network measurements. **(ii)**

Hierarchical learning at the Internet: investigates the main properties of hierarchical learning approaches, through the application of divide & conquer paradigms to solve complex learning tasks, breaking them up into smaller components. The Internet itself has a highly hierarchical structure, thus the synergies and parallels between learning with hierarchical data and the underlying hierarchical networks is an appealing topic to explore. **(iii) Multi-task learning:** lastly, I put the emphasis on learning from multiple, heterogeneous networking data, aiming to generate knowledge from the large amount of data available in large scale networks. I propose to leverage the data contained in multiple related tasks to improve overall learning generalization, by solving multiple learning tasks at the same time, while exploiting commonalities across tasks.

Finally, **AI Defined Networking** has as main objective to make of AI/ML and the knowledge distributively generated at the Internet scale a driving force for flexible network management and operation. With the recent developments and rising deployment of SDN at the large scale [26], there is now a bright new path to couple this flexibility with the knowledge generated by AI/ML. Here I propose to **(i) revamp the KP (Knowledge Plane)** concepts [12], focusing in particular on the KP architecture in terms of availability and accessibility of the knowledge (trained-models, building-blocks) generated across multiple networks. I propose to investigate different concepts and architectural designs to make this knowledge flowing through the Internet, the same way routing information and content propagate in current networks. How to store and distribute such knowledge brings me to a novel proposal, that of **(ii) Knowledge Delivery Networks (KDNs)**. A KDN replicates the same concepts of current Content Delivery Networks (CDNs), but takes into account the potential and characteristics of the particular content, similarly to the Information-Centric Networking (ICN) paradigm [25]; using the techniques and results obtained through effective, Internet learning, a KDN could build, maintain and merge different pieces of *knowledge-blocks* to improve learning and replicate lessons learned. The KDN concept would eventually lead to a knowledge transfer system, where different models learned in specific environments and conditions could be re-used, to improve new learning steps. Similar ideas have been recently proposed [36], relying on Cloud repositories. A KDN overlay network could be initially realized through well established CDN networks. **(iii) AIDN meets SDN:** the ultimate integration of the concepts behind Internet Learning, the foundations of the KP, and the realization of KDN networks withing a SDN architecture, would lead to the next generation of AI/ML-defined (driven and capable) networks.

V. Conclusions

The ultimate goal of this positioning paper is to contribute to strengthening the future research on the (re)-emergent field of AI4NETS, by pinpointing some of the fundamental challenges it faces. I acknowledge that the discussion I present in this paper is by no means fully exhaustive and final, and I am sure many aspects were left aside; however, I do hope it will motivate further discussion on the limitations we are facing in AI4NETS as a research community, to better target the next wave of AI/ML-based solutions in the practice.

References

[1] *Journal on Machine Learning, Special Issue on Inductive Transfer*, 28(1), 1997.

[2] M. A. Alsheikh, S. Lin, D. Niyato, and H. Tan. Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications. *IEEE Communications Surveys Tutorials*, 16(4):1996–2018, Fourthquarter 2014.

[3] Y. Bengio, A. Courville, and P. Vincent. Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1798–1828, Aug 2013.

[4] C. M. Bishop. *Pattern Recognition and Machine Learning, 5th Edition*. Springer, 2007.

[5] M. Bkassiny, Y. Li, and S. K. Jayaweera. A Survey on Machine-Learning Techniques in Cognitive Radios. *IEEE Communications Surveys Tutorials*, 15(3):1136–1159, Third 2013.

[6] A. Blenk, P. Kalmbach, W. Kellerer, and S. Schmid. O'Zapft is: Tap Your Network Algorithm's Big Data! In *Proceedings of Big-DAMA'17*, pages 19–24, 2017. ACM.

[7] R. Boutaba, M. A. Salahuddin, N. Limam, S. Ayoubi, N. Shahriar, F. Estrada-Solano, and O. M. Caicedo. A Comprehensive Survey on Machine Learning for Networking: Evolution, Applications and Research Opportunities. *Journal of Internet Services and Applications*, 9(1):16, Jun 2018.

[8] J. V. Bouvrie. *Hierarchical Learning: Theory with Applications in Speech and Vision*. MIT Press, 2009.

[9] M. L. Brodie, editor. *Making Databases Work: The Pragmatic Wisdom of Michael Stonebraker*. Association for Computing Machinery and Morgan Claypool, New York, NY, USA, 2019.

[10] A. L. Buczak and E. Guven. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys Tutorials*, 18(2):1153–1176, Secondquarter 2016.

[11] R. Caruana. Multitask Learning. *Machine Learning*, 28(1):41–75, Jul 1997.

[12] D. D. Clark, C. Partridge, J. C. Ramming, and J. T. Wroclawski. A Knowledge Plane for the Internet. In *Proceedings of SIGCOMM '03*, 2003.

[13] A. Crotty, A. Galakatos, and T. Kraska. Tupleware: Distributed machine learning on small clusters. *IEEE Data Eng. Bull.*, 37(3):63–76, 2014.

[14] J. Crowcroft. Internet Failures: an Emergent Sea of Complex Systems and Critical Design Errors? *The Computer Journal*, 53(10):1752–1757, Dec 2010.

[15] J. Dean, G. Corrado, R. Monga, K. Chen, M. Devin, M. Mao, M. A. Ranzato, A. Senior, P. Tucker, K. Yang, Q. V. Le, and A. Y. Ng. Large Scale Distributed Deep Networks. In F. Pereira, C. J. C. Burges, and L. Bottou editors, *Advances in Neural Information Processing Systems 25*, pages 1223–1231. 2012.

[16] Z. M. Fadlullah, F. Tang, B. Mao, N. Kato, O. Akashi, T. Inoue, and K. Mizutani. State-of-the-Art Deep Learning: Evolving Machine Intelligence Toward Tomorrow's Intelligent Network Traffic Control Systems. *IEEE Communications Surveys Tutorials*, 19(4):2432–2455, Fourthquarter 2017.

[17] A. Fischer and C. Igel. Training Restricted Boltzmann Machines: an Introduction. *Pattern Recognition*, 47(1), 2014.

[18] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau. An Introduction to Deep Reinforcement Learning. *CoRR*, abs/1811.12560, 2018.

[19] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.

[20] I. Goodfellow, P. McDaniel, and N. Papernot. Making Machine Learning Robust Against Adversarial Inputs. *Commun. ACM*, 61(7):56–66, June 2018.

[21] W. Guo, D. Mu, J. Xu, P. Su, G. Wang, and X. Xing. LEMNA: Explaining Deep Learning Based Security Applications. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, CCS '18, pages 364–379, New York, NY, USA, 2018. ACM.

[22] G. E. Hinton and R. R. Salakhutdinov. Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786):504–507, July 2006.

[23] J. Jiang, V. Sekar, I. Stoica, and H. Zhang. Unleashing the Potential of Data-Driven Networking. In *Communication Systems and Networks - COMSNETS*, 2017.

[24] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza. A Survey of Machine Learning Techniques Applied to Self-Organizing Cellular Networks. *IEEE Communications Surveys Tutorials*, 19(4):2392–2431, 2017.

[25] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica. A data-oriented (and beyond) network architecture. In *Proceedings of SIGCOMM '07*, 2007.

[26] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig. Software-Defined Networking: A Comprehensive Survey. *Proceedings of the IEEE*, 103(1):14–76, Jan 2015.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25*, pages 1097–1105. 2012.

[28] Y. LeCun, Y. Bengio, and G. E. Hinton. Deep Learning. *Nature*, 521(7553):436–444, 2015.

[29] C. Lemke, M. Budka, and B. Gabrys. Metalearning: a Survey of Trends and Technologies. *Artificial Intelligence Review*, 44(1):117–130, Jun 2015.

[30] S. M. Lundberg and S.-I. Lee. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30*, pages 4765–4774, 2017.

[31] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y. Liang, and D. I. Kim. Applications of Deep Reinforcement Learning in Communications and Networking: a Survey. *CoRR*, abs/1810.07862, 2018.

[32] Q. Mahmoud. *Cognitive Networks: Towards Self-Aware Networks*. Wiley-Interscience, New York, NY, USA, 2007.

[33] H. Mao, M. Alizadeh, I. Menache, and S. Kandula. Resource Management with Deep Reinforcement Learning. In *Proceedings of the 15th ACM Workshop on Hot Topics in Networks*, HotNets '16, pages 50–56, New York, NY, USA, 2016. ACM.

[34] H. Mao, R. Netravali, and M. Alizadeh. Neural Adaptive Video Streaming with Pensieve. In *Proceedings of SIGCOMM '17*, 2017.

[35] G. Marín, P. Casas, and G. Capdehourat. DeepSec Meets RawPower - Deep Learning for Detection of Network Attacks Using Raw Representations. *ACM SIGMETRICS Perform. Eval. Rev.*, 46(3):147–150, Jan. 2019.

[36] M. Mehari. *Performance Optimization and Modelling of Complex Wireless Networks Using Surrogate Models*. Ghent University, Faculty of Engineering and Architecture, Gent, Belgium, 2018.

[37] A. Mestres, A. Rodriguez-Natal, J. Carner, P. Barlet-Ros, E. Alarcón, M. Solé, V. Muntés-Mulero, D. Meyer, S. Barkai, M. J. Hibbett, G. Estrada, K. Ma'ruf, F. Coras, V. Ermagan, H. Latapie, C. Cassar, J. Evans, F. Maino, J. Walrand, and A. Cabellos. Knowledge-Defined Networking. *SIGCOMM Comput. Commun. Rev.*, 47(3), 2017.

[38] P. Mulinka, S. Wassermann, G. Marín, and P. Casas. Remember the Good, Forget the Bad, do it Fast: Continuous Learning over Streaming Data. In *Advances in Neural Information Processing Systems 32, Continual Learning Workshop*. 2018.

[39] T. T. T. Nguyen and G. Armitage. A Survey of Techniques for Internet Traffic Classification using Machine Learning. *IEEE Communications Surveys Tutorials*, 10(4):56–76, 2008.

[40] G. I. Parisi, R. Kemker, J. L. Part, C. Kanan, and S. Wermter. Continual Lifelong Learning with Neural Networks: A Review. *CoRR*, abs/1802.07569, 2018.

[41] B. J. Radford, L. M. Apolonio, A. J. Trias, and J. A. Simpson. Network Traffic Anomaly Detection Using Recurrent Neural Networks. *CoRR*, abs/1803.10769, 2018.

[42] M. T. Ribeiro, S. Singh, and C. Guestrin. Why Should I Trust You?: Explaining the Predictions of Any Classifier. In *Proceedings of the 22Nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2016.

[43] S. J. Russell and P. Norvig. *Artificial Intelligence - A Modern Approach*. Pearson Education, 2010.

[44] J. Saxe, R. Harang, C. Wild, and H. Sanders. A Deep Learning Approach to Fast, Format-Agnostic Detection of Malicious Web Content. In *2018 IEEE Security and Privacy Workshops (SPW)*, pages 8–14, 2018.

[45] J. Schmidhuber. Deep Learning in Neural Networks: an Overview. *Neural Networks*, 61:85–117, 2015.

[46] R. Sommer and V. Paxson. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. In *Proceedings of the 2010 IEEE Symposium on Security and Privacy*, SP '10, pages 305–316, 2010.

[47] R. S. Sutton and A. G. Barto. *Reinforcement learning - an Introduction*. Adaptive computation and machine learning. MIT Press, 1998.

[48] R. W. Thomas, L. A. DaSilva, and A. B. MacKenzie. Cognitive networks. In *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005.*, pages 352–360, Nov 2005.

[49] E. Zamanian, C. Binnig, T. Kraska, and T. Harris. The End of a Myth: Distributed Transaction Can Scale. *PVLDB*, 10(6):685–696, 2017.

[50] J. Zhao, S. Shetty, and J. W. Pan. Feature-based Transfer Learning for Network Security. In *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, pages 17–22, 2017.