Deploying Artificial Intelligence in the Wireless Infrastructure: the Challenges Ahead

Miguel Ángel Vázquez
Centre Tecnològic de Telecomunicacions de Catalunya
Castelldefels, Spain
mavazquez@cttc.es

Jean Paul Pallois
Huawei Technologies
Paris, France
jp.pallois@huawei.com

Merouane Debbah
Huawei Research Centre
Paris, France
merouane.debbah@centralesupelec.fr

Christos Masouros
University College London
London, United Kingdom
c.masouros@ucl.ac.uk

Tony Kenyon
University College London
London, United Kingdom
a.kenyon@ucl.ac.uk

Yansha Deng
King's College London
London, United Kingdom
yansha.deng@kcl.ac.uk

Fisseha Mekuria
Council for Scientific and Industrial Research
Pretoria, South Africa
FMekuria@csir.co.za

Ana Pérez-Neira
Centre Tecnològic de Telecomunicacions de Catalunya and Univeritat Politècnica de Catalunya
Castelldefels, Spain
aperez@cttc.es

Javan Erfanian
Bell Mobility
Ontario, Canada
javan.erfanian@bell.ca

***Abstract*— The adoption of artificial intelligence (AI) techniques entails a substantial change in the wireless ecosystem where data as well as their owners become crucial. As a result, the roll out of AI techniques in wireless systems raises a plethora of questions. In this context, we describe the challenges observed by the wireless stakeholders when deploying AI. Furthermore, we introduce the recent discussion in field of ethics that appear when managing wireless communications data.**

Keywords—Artificial Intelligence, Machine Learning, Wireless Infrastructure,

I. Introduction

Telecom operators see AI as almost the unavoidable technology enabling to maintain or even reduce the operational expenditures (OPEX) significantly while delivering higher quality-of-service (QoS) to the end-users. Communication service providers will be looking at AI for various reasons. On top of enhancing their business ability, the adoption of AI will be driven by the increased complexity of the network.

Indeed, with the advent of 5G, networks will be handling more spectrum, more and varied bandwidths, additional radio technologies, dealing with lower latencies enabling to reach new business territories like on-line tactile applications while benefiting from always more computing power. AI has the power to change the networks of the future from reactive networks, to predictive and finally proactive networks. Insights from the deep learning AI systems using the huge amount of data generated by the complex wireless systems of the future, will be able to support effective utilization of spectrum & radio resources, self-optimization of network operations and insight based QoS provisioning, thereby benefitting the wide spread application of AI in telecoms. However, AI is not seen as the grail technology by everyone. A few factors, not necessarily negligible, may slow down the adoption of AI by the industry. The success of AI in wireless networks of the future therefore hinges, in developing solutions that will address these challenges starting from the design of the network layers and up to the service layers.

In this paper, we describe those factors in the next Section. In addition, in Section III we describe the first issues raised by the stakeholders regarding the proper use of the wireless data. Section IV concludes.

II. AI Challenges

The first reason may rely on the difficulty to fully assess the real benefits AI can bring in daily work despite the fact the market feels it is the way to go. Another reason is also a human, emotional reason: people in the industry may perceive AI as a potential threat to jobs. Right or wrong, this is a reality, AI proponents will have to deal with. But even if the AI advocates pass these barriers, the need to adapt existing network operations processes will be a heavy, cumbersome and tedious task and will need support from the top management in organizations adopting AI. An additional obstacle will be the expected transparency of AI. Indeed, delivering AI as a black box may be enough in some cases for basic actions (e.g. like how to optimize parameter settings during a roll-out phase) but would not be acceptable during certain or healing phases. Operators who could not be in the position to explain what happened and why it happened could be reliable to liabilities. Here in lies the next obstacle, which is the training data set accuracy and inclusiveness, and might result in a biased decision making which can be detrimental to the company and its customer base. This is why explainable AI will be sooner or later necessary. All these factors will need, undoubtedly, to be addressed from a successful adoption of AI by the market and the telecommunications industry.

Importantly, in the dynamic and real time wireless transmission scenarios of the future, the training overheads for learning based AI solutions may pose a particular challenge. As an example, typical learning tools available have been developed for applications such as computer vision, speech recognition, or natural language processing, where training time and overheads are not a key factor. In a dynamic wireless environment where the transceivers will need to adopt to changing channels, hardware responses, or event dynamic link connectivity, the overheads of re-training

978-1-7281-3627-1/19/$31.00 ©2019 IEEE

the system are paramount. Solutions inspired by layered and predictive training such as incremental or reinforcement learning will inevitably become key in addressing such environments.

The widespread adoption of AI pose severe challenges to the speed and power consumption of existing network devices and wireless systems. A recent Nature editorial put it succinctly by asking the question “Does AI have a hardware problem?”, the author’s answer being a definitive “yes” [1]. Present day von Neumann computing architectures require constant shuffling of data between storage and CPU, providing a critical performance bottleneck. Most clock cycles are wasted in moving data rather than computing, while physical separation of memory and processing builds in latency. AI systems, which strive to mimic some functions of the brain, have a significant power cost. While the human brain expends ~20W, large AI systems can consume tens of kW. Scaling up becomes prohibitive; for example, a simulation of a neural network approaching the complexity of the human brain (10^{10} neurons and 10^{14} synapses), running on the Lawrence Livermore Sequoia supercomputer, consumed 7.9MW [2]. Optimized graphic processing units (GPUs) and tensor processing units (TPUs) offer significant benefit, but still consume orders of magnitude more power than biology when performing similar tasks, and do not offer the fault and noise tolerance of biological systems.

The international technology roadmap for semiconductors (ITRS) and the United States department of energy office of science recognize this as a key challenge, stating in 2015: “*Well-supported predictions ... indicate that conventional approaches to computation will hit a wall in the next 10 years ... Novel approaches & new concepts are needed…*” [3].

As we begin to deploy AI systems both at the heart of communication systems and as edge computing elements for the IoT, such power consumption issues will severely limit the possibility of integrating AI and communication unless significant effort is directed to more efficient hardware solutions. While these will undoubtedly include further optimization of coprocessors (GPUs, TPUs), more ambitious solutions, such as non-von Neumann approaches, are required in the long term.

III. Ethics

In recent months and weeks, some important studies have been issued on the topic of ethics for handling data. One of the first was issued by the French commission nationale de l’informatique et des libertés (CNIL) in 2017. Amongst the six policy recommendations elaborated by the CNIL, the recommendation #2 is directly related to the usage of AI in wireless networks: “*Making algorithmic systems comprehensible by strengthening existing rights and by rethinking mediation with users*”.

At the time of writing, the latest report on ethics has been issued by the European Union in April 2019. The EU has released a report including recommendations and requirements to achieve trustworthy AI: “*AI Ethics Guidelines for Trustworthy AI – April 2019*”. In this comprehensive report, one requirement on “privacy and data governance” relates to the point raised above. A few key additional needs apply directly to wireless networks as well. In light of the mentioned report, to realize trustworthy AI-based wireless networks the following requirements will have to be fulfilled: 1) Technical robustness and safety, 2) Privacy and data governance, 3) Transparency and 4) Accountability. Both reports, the one from the French CNIL and the one from the European Union, highlight the need of explicability, transparency and accountability. These requirements are paramount with the advent of 5G. 5G will enable new businesses and address vertical ones such as the automotive, transport and logistics or health industries for instance. Accountability and liability will be of the utmost importance. In case problems arise, the 5G network provider will have to explain what happened, why it happened and when the decision was taken. As a consequence, network providers will need to have access to logs, traces and understand the type of algorithms used while having access to the details to the logics that conducted to certain actions. However, it has to be noted that not all actions empowered by AI taken by the network will be needed to be fully explainable. For example, fault monitoring and analysis enhanced by AI to reveal concealed anomalies may not need to deliver a full detailed explanation of the intelligence used. On the other hand, an algorithm guaranteeing a very low latency delivered by a 5G network instantiation (5G slice) for health care and remote surgery for instance, will need to deliver details in case of any issue arise due to legal liabilities and the fact that AI does not work in a lawless world. It is the same for connected cars or intelligence transportation. As soon as legal responsibilities are at stake, the network powered by AI must be a trustworthy AI based network and be able to explain its actions.

Next generation wireless networks will be continuing to play a major transformative role and deliver technical and social benefits to the society. AI will enhance the way wireless networks behave and make them the central nerve system of the future telecoms economy. But the success of it will be relying on secure, trustworthy application of AI in future networks guaranteeing the privacy of everyone.

IV. Conclusions

AI algorithms and technologies, promise a huge potential, for improved network resource utilization, reduced cost and generate new business insights from the huge amount of data generated in future wireless communications systems. At the same time, the paper highlights that the proliferation and pervasive deployment of AI will generate potential entanglement in the form of increased power consumption, latency and QoS which require due attention. Finally, the ethical application of AI with respect to the stakeholders of communications network data-owners, service providers, regulators and end-users, need to be established, as a key enabler for successful deployment of AI in communications infrastructure and systems.

Acknowledgment

The authors would like to thank all members of the AI working group of World Wireless Research Forum for their fruitful discussions.

References

[1] Does AI have a hardware problem?”, Nature Electronics (Editorial) vol. 1 no. 4, 205, 2018

[2] T. Wong et al, IBM Res. Rep. RJ10502 (2012)

[3] I.K. Schuller & R. Stevens, “Neuromorphic computing: from materials to systems architectures”. Round Table Rep

---

This work has received funding from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economia y Competitividad) under project TEC2017-90093-C3-1-R and from the Catalan Government (2017 SGR 1479). Part of this work was funded by the EPSRC project number EP/S028455/1