SCTE·ISBE CABLE-TEC EXPO 2019 | NEW ORLEANS, LA SEPT. 30-OCT. 3

2019 Fall Technical Forum
SCTE·ISBE • NCTA • CABLELABS

Can Future Networks Survive Without Artificial Intelligence?

A Comprehensive Overview of

Frameworks and Applications of AI

A Technical Paper prepared for SCTE•ISBE by

Dr. Claudio Righetti
Chief Scientist
Telecom Argentina S.A.
Agüero 2392, Buenos Aires, Argentina
Phone: +5411 5530 4468
crighetti@teco.com.ar

Emilia Gibellini
Scientist
Telecom Argentina S.A.
egibellini@teco.com.ar

Carlos Germán Carreño Romano
Scientist
Telecom Argentina S.A.
caromano@teco.com.ar

Gabriel Carro
VP Engineer and R&D
Telecom Argentina S.A.
gcarro@teco.com.ar

Table of Contents

List of Figures

Abstract

In the last two years, Artificial Intelligence (AI) has become a trending topic in the main conferences and international exhibitions. One of the first questions that arises is to refer to the operators and vendors when they use the term AI. These days it really seems that everything gets better with AI, then another important point is hype vs reality. Certainly, there has been a lot of activity in trying to use AI to improve cable network operation and customer experience. On the other hand, future networks will generate a volume of data with exponential growth, and this poses a great challenge. We present in this technical paper a framework of AI technologies, how operators can extract the value from data and gain new insights and efficiencies using artificial intelligence, where to start (the right points) in our networks to apply AI and what human skills will be needed. Finally, we present our biggest challenge: our AI and Data Strategy that will drive our transformation from an MSO and MNO to a Digital Service Provider (DSP).

Content

1 Introduction

In our networks and services, AI has the potential to change, the way we operate, and to become the foundation of the transformation that leads to the fourth industrial revolution. But this requires hard work, a long-term commitment, and a deep cultural change.

Telecom Argentina is currently in a process of integration of its MSO and MNO networks. To accompany the fourth industrial revolution, we need to become a Digital Service Provider (DSP). Telecom is going to a Multiservice Convergent Network and Multi devices approach, where the Client / User can consume their own and third party services (platforms) from any device and connected to any of our access networks.

A digital service is defined by The European Union (EU) Agency for Cybersecurity, more specifically by the Network and Information Security (NIS) directive 2015/1535 as “any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services”. According to the NIS directive, DSPs are limited to only three types of services: cloud, online market places and search engines [1].

The term Digital Service Provider applies to any company that distributes media online. In the case of telcos, it is an organization that has moved on from offering core, traditional telecom services, to providing mobile broadband access, services, content and apps, all sold directly from the device [2].

The DSP is not merely a dumb pipe offering shared access to a common utility; it is an online, real-time business that deals with countless transactions every day, managing high volumes of data traffic and multiple devices per user, and often multiple users per account. The mobile and fixed landscape has changed dramatically and CSP’s are fine-tuning their businesses, and their network infrastructure, to cater for the digital needs of the data-hungry customer.

1.1 Motivations and Overview

The term AI has recently become a buzzword and entered in a sort of semantic satiety similar to what happened a few years ago with the term Big Data. But what exactly is AI? What is it used for? Who invests in it? Can future networks survive without it?

All Operations Support Systems (OSS) in our current and future networks generate a huge amount of data. How can we generate a culture of innovation driven by data? What is the value of the data?

The final aim of all of these efforts is to be able to offer our services in an adequate way for our next generations of clients. They are nowadays putting the requirements in the market and driving the evolution of technologies. Ultimately our clients do not buy technologies, they buy services.

All these services will be enabled by technologies such as 10G, 5G, SDN / NFV, Holographics Displays etcetera. Operating, managing and provisioning future services with automation processes becomes essential.

Operating, managing, provisioning future services with automation processes must be essential. If we want a complete automation, we will need AI technology and data analytics tools.

This technical paper proposes a comprehensive overview of frameworks and applications of AI to network's design, management and operation. We introduce a distributed cognitive system that permeates the network, that is called the knowledge plane. This paper is organized as follows. Following this section we define AI and Data Analytics. In section 2, we expose some applications of AI and Data Analytics to communications, networks and services. Section 3 exposes the technical and theoretical framework, we briefly describe some ML and non-ML algorithms that have found useful and how we are applying them, and we introduce the concept of the knowledge plane. The last section, which is section 4, outlines the key challenges in our path.

1.2 Artificial Intelligence

In this work we define our scope of AI, which is much more related to the current Machine Learning (ML) framework, we propose some technical and theoretical issues of AI, and an overview of applications to networks. In the next sections we hope to answer some of the starter questions.

The concept of AI is not new. It was already established in the 1950's as machine intelligence that is capable to process, analyze, and react to input and changing situations by itself. We can say that today AI is a general term that describes multiple technologies, none of which fit completely to the original definition of AI.

According to our chosen bibliography, intelligence is the capacity to do the right thing at the right time, in a context where doing nothing (or making no change in behavior) would be worse. Intelligence then requires the capacity to perceive contexts for action, the capacity to act, and the capacity to associate contexts to actions. AI, by convention, describes (typically digital) artifacts that demonstrate any of these capacities [3].

1.2.1 A brief history

In the 1940's and 1950's, scientists in the fields of Mathematics, Engineering and Computer Science explored the possibilities of artificial brains and tried to define the intelligence of a machine. In 1950, Alan Turing presented a test known today as the Turing test, which defined the concept of Machine Intelligence [4]. John McCarthy is credited with the first definition of AI as *the science and engineering of making intelligent machines*, during a workshop at Dartmouth College in 1956 [5]; Russell and Norvig present a classification of the available definitions of AI in two senses: one based on the function expected to be performed (comparing processes/reasoning of the machine versus the outcome/behavior that it exhibits) and the other about the metrics used for assessing the success of AI (human performance versus an ideal standard of rationality) [6], and in their seminal 1995 book, they give it an arresting description, *agents that* *receive percepts from the environment and take actions that affect that environment*. In Figure 1, we present the evolution and stages of AI.

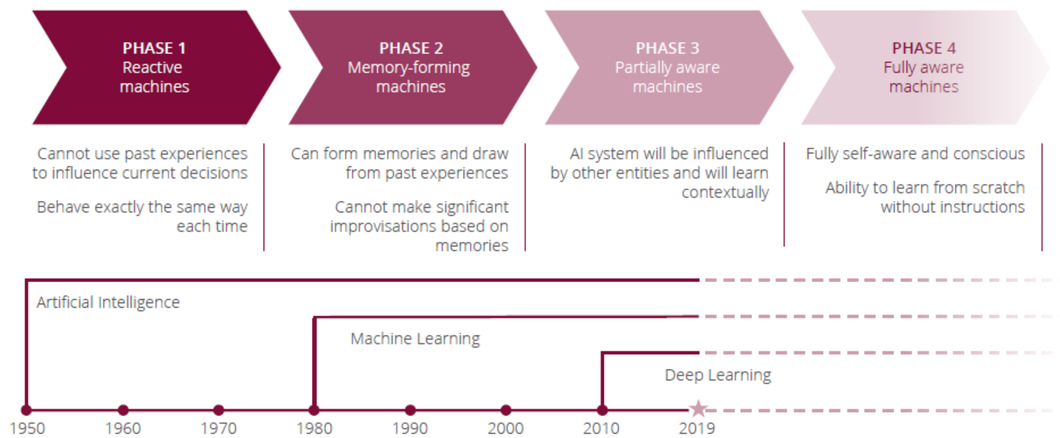

Figure 1 - The evolution of Artificial Intelligence.
Source: Statista, Nvidia.

In 1959 Machine Learning (ML) arises, Arthur Samuel defines it as *the field of study that gives computers the ability to learn without being explicitly programmed* [7]. So, the word learning in this sense is used by analogy with the learning process in animals rather than in humans.

As we see the term AI is, in broad sense, the human intelligence process replicated by a machine, which is an auto programmable entity. People think in robots, in the Turing test, a person asking a robot and being answered by an entity which could be a human itself, something like a chatbot but more realistic. Critics go deeply through the philosophical plane.

In our opinion, this meaning will not be possible at least in the next few years except in movies, so we prefer to write about AI in a narrow sense. It is interpreted as a subset of programming techniques based on statistics, sometimes referred to as ML. In order to be more specific, our approach comprehends the collection of data, polling of the network, generation of useful models, information and algorithms based on business rules. From a statistical outlook, it is an expansion of the classical statistical methods which includes the conception supervised, unsupervised and reinforcement methods. From a computer science point of view, AI differs from classical programming which basically is rules + data, in the idea of rules learned from the data.

In recent years, the progress in AI has been driven mainly by the generation and availability of huge amounts of data and augmented computing power.

1.3 Data Analytics

In the last two years, 90% of the world's total data load has been generated, yet only 1% has been processed [8]. With data volumes set to increase exponentially (Figure 2) over the next decade, significant challenges are waiting on our journey.

Companies who can extract the value from data and gain new insights and efficiencies using AI and Data Analtytics could boost the fourth industrial revolution [9]. *Data is going to introduce social and economic changes that we see perhaps once or twice in a century*, said ex-CEO of Intel, Brian Krzanich, in his keynote speech at CES 2018.

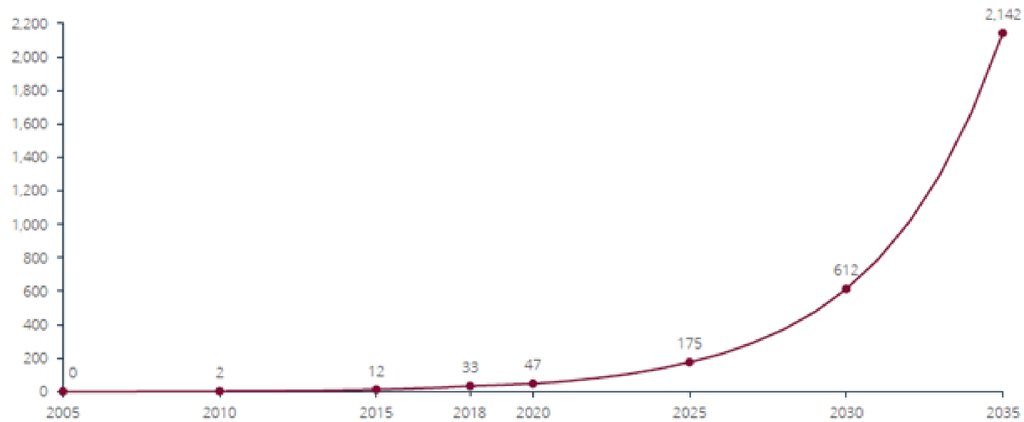

Figure 2 - Worldwide amount of data created per year in zettabytes.
Source: Statista.

One of the main differences between Data Analytics and ML is that the former extracts value from data by applying exploratory techniques while the latter implies a certain level of inference.

2 AI and Data Analytics in Communications, Networks and Services

There are several organizations and working groups proposing frameworks, standards and how to apply the AI to the industry of communications, networks and services. Below we detail those in which Telecom Argentina is participating or is in consultation with:

- SCTE & CableLabs AI/ML Working Groups
- Telecom Infra Project - AI/ML Working Group
- TM Forum - AI & Data Analitycs
- ITU-T Study Group 13
- ITU-T FG ML5G (Est. in 11/2017) Studying network architectures, use cases, and data formats for the adoption of machine learning methods in 5G and future networks.
- ETSI ISG ENI (Experiential Network Intelligence) (Est. in 2/2017) Defining a cognitive network management architecture based on AI methods andcontext-aware policies; five deliverables have already been released

At the TM Forum [10] presentations and discussion panels, possible uses and applications in the telecommunications business were presented. Among the advantages of the use of ML, we can mention:

- Fast and automatic analysis of large volumes of data that are becoming more and more complex.
- Getting faster and more accurate results that allow to make reliable and repeatable decisions.
- Focus on behavioral analysis to detect and predict possible anomalous events at an early stage.
- Automate real-time analysis in the orchestration of end-to-end services in a virtualized world.
- Identification and mitigation of security threats in services through predictive analytics and ML to detect attacks that escape traditional preventative static defenses.

- Prediction of churn. Unlike traditional strategies, ML allows a multi-class classification of our clients, for example to predict whether they belong to a low, medium or high-risk class.
- Support for automation and management of network orchestration and traceability of end-to-end transactions across the network and OSS/BSS environment.

For the cable industry in particular, [11] provides an overview of ML algorithms, and how their potential applications could be applied:

- Software Defined Networks (SDN) Routing
- Profile Management on DOCSIS 3.1 cable modems [12]
- Proactive Network Maintenance (PNM): for DOCSIS
- HFC´s Network Health KPI

Some applications are being implemented in:

- Internet Traffic Characterization
- Network Traffic Engineering
- Wi-Fi Proactive Network Maintenance (PNM)

2.1 Definitions

We have taken some definitions from those work groups and we started to build our own AI and Data Analytics technologies application framework in order to apply it to the Telecom Argentina’s networks and services. We will mention some of the definitions from the framework.

- **Data Analytics**: monitoring data to look for patterns and anomalies (without applying intelligence) and applying those patterns towards effective decision making.
- **Artificial Intelligence**: the development of computer systems capable of performing tasks that normally require human intelligence; this includes visual perception, speech recognition, decision-making, and translation between languages.
- **Machine learning**: a type of AI that gives machines the ability to learn automatically and improve from experience without being explicitly programmed.
- **Deep learning**: takes machine learning further by processing information in layers, where the result or output from one layer becomes the input for the next.
- **Automation:** within MSOs and MNOs, this means automation of processes that were previously carried out by people; AI is an enabling technology that may (or may not) help with the process of automation.
- **Cognitive computing:** like AI, cognitive computing is based on the ability of machines to sense, reason, act and adapt based on learned experience, but whereas AI acts on its analysis to complete a task, cognitive computing provides the information to help a person decide.
- **SON** (Self-Organizing Networks): a technology for automating the planning, configuration, management, optimization and healing of mobile radio access networks; it was developed by 3GPP and is sometimes conflated with AI.
- **Explainable AI:** explainability sits at the intersection of transparency (consumers have the right to have decisions affecting them explained in understandable terms), causality (it is expected of the algorithms to provide not only inferences but also explanations), bias (the absence of bias should be guaranteed), fairness (it should be verified that decisions made by AI are fair) and safety (reliability of AI systems) [13].

Conceptually speaking, it is clear from Figure 3 that machine learning is a subset of AI, which includes deep learning algorithms.

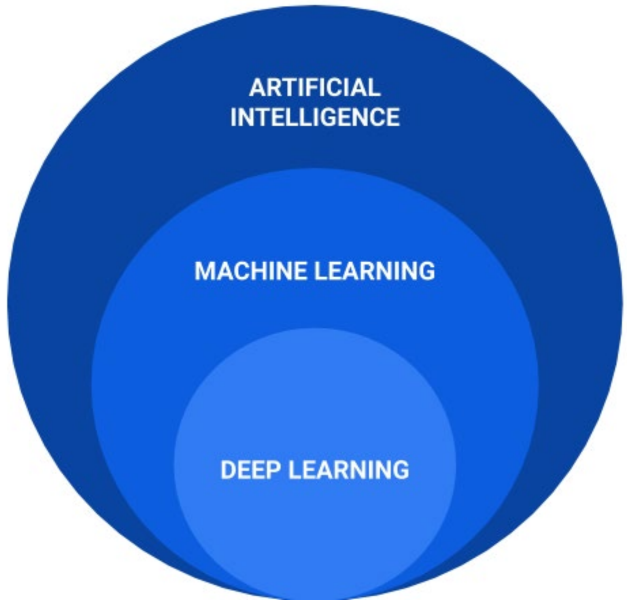

Figure 3 - Map of AI

Regarding this last point, and according to [14], [15] we know that many machine learning algorithms have been labeled “black box” models because of their inscrutable inner-workings. What makes these models accurate is what makes their results difficult to interpret and understand; they are very complex. So, even when some abstraction or transformation models can be explainable, not always they are auditable. The discussion about audit AI is still open [16].

2.2 Applications

Several applications from different fields of knowledge are being hyped during last few years. As shown in Figure 4, some of them include the transport and logistics industry, healthcare systems, agriculture, semiconductors industry, among others. In this document we will focus on network and service-related applications.

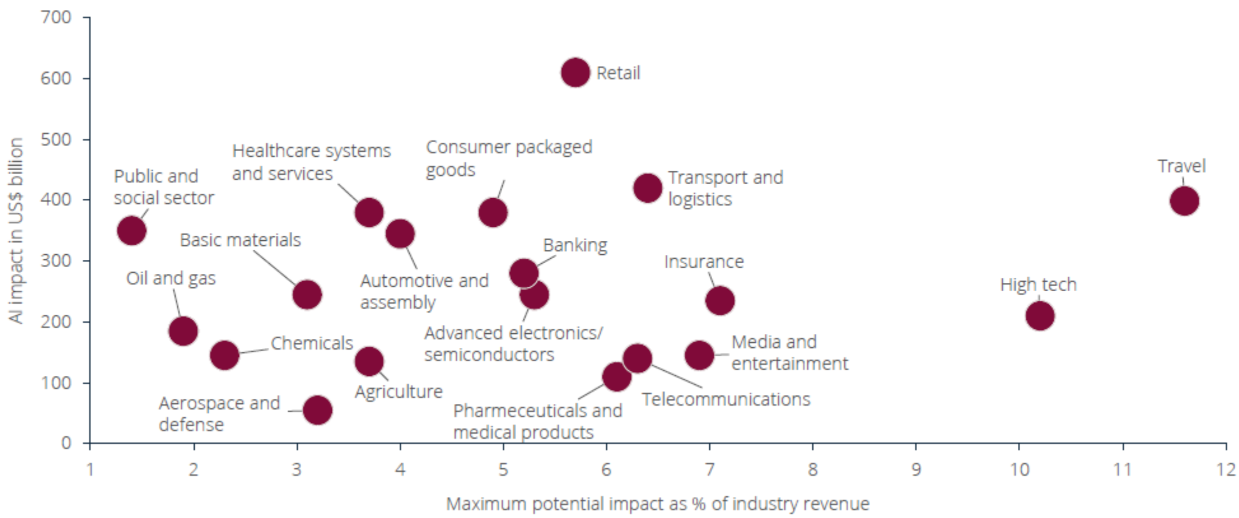

Figure 4 - Potential annual impact of AI technologies in global industries by 2025.
Source: McKinsey Global Institute, Statista.

From a global or top-level perspective, we will mention and describe some of the current and next few years applications of AI applied to networks and services, which we consider of potential interest in these industries. Some of them are related to current deployments and some of them are being tested in lab.

- **Management**

Cable Television industry, Telco and TechCo in general have some technical concerns at the management level. Some of them are the specific technology election for traffic classification and prioritization, the licensing of the spectrum, the outside plant, the attractiveness of the website, the commercial and technological strategy.

Dealing with the mass of users, devices and applications implies certain requirements for the network in order to service data. AI must be one of the strategic management tools for different purposes. Depending on the scenario, it could be used for reducing costs, customizing the user preferences, launching dynamic offers to the market, etcetera.

- **Maintenance**

Maintenance efforts, including corrective, predictive and proactive maintenance are very important. Nowadays Proactive Network Maintenance (PNM) work groups show a good picture about the need of ML algorithms for the classification of network impairments. Power control schemes could be improved also with tools of monitoring and analytics that combine not only hardware and systems raw data but also user experience information coming from, for example, social networks.

In the telco industry there is a trend of monitoring the user experience through specific teams associated with the focus on the service, and not only on the network. Despite the effort to change the vision, this is still considered as a way of corrective maintenance because it reacts when something is wrong. But there are also new ideas, for example digital twins, which could be used not only for troubleshooting but also for simulation of network changes. This approach currently involves tools for testing the effectiveness of a change at low costs through A/B testing. We will talk about these technologies in the rest of the document.

- **Engineering and architecture**

In 5G, big data and analytics are leveraged to extract massive patterns, especially at the physical (PHY) and medium access control (MAC) layers and enable self-organizing operations. Artificial neural networks (ANN) can be used to redefine communication networks, solving a number of non-trivial design problems at runtime and across layers for cognitive link adaptation, resource scheduling, signal classification, and carrier sensing/collision detection, among others [17].

For fixed networks there are several step by step changes within the SDN/NFV roadmap, which implies and make possible the application of algorithms for network deployment at the orchestration level. Machine Learning techniques are interesting in the administration of cloud platforms, particularly when dynamic scaling of capacity is a requirement within the applications. We will cover some notions about technical concerns of these technologies.

The design and optimization of wireless communication systems are becoming more and more challenging, due to the extreme key performance indicators (KPIs) for user experience, efficiency, performance and complex network environments. AI, which can exploit the increasingly massive datasets available from wireless systems, can be used to solve complex and previously intractable problems.

Many problems in wireless communication systems, such as decision making, resource optimization, and network management, can be cast in a form that is suitable to be solved by AI techniques. Based on this observation, it is essential to study when and how to apply AI technology to improve the performance of future wireless communication systems.

2.2.1 Traffic classification

Characterization of network traffic has been proved to be useful to many applications, such as network engineering and management, quality of service (QoS) assurance, distribution of tiered services, automated re-allocation of network resources, security, and even censorship. It provides data on subscriber's usage, which enables the application of data mining to support business goals. Furthermore, governments expect the CSPs to provide the facilities for lawful interception, for which traffic classification is key

The classic approach, which consists on the observation of the port number, is not effective since some applications use well-known port numbers to disguise their traffic or use unpredictable port numbers. Data packet inspection (DPI) techniques base their strategy on the search of keywords in the data packets' headers, which can be expensive in terms of computational effort and infeasible in case the data is encrypted. Moreover, governments may impose privacy regulations constraining the ability of third parties to lawfully inspect payloads at all. In consequence, a new generation of methods, based on machine learning (ML), have been proposed [18].

For the application of ML algorithms, the analysis unit is the flow, which consists of a succession of IP packets that have the same protocol type, source address and destination address. Most approaches consist of classifying the flows according to their statistical properties (distribution of flow duration, flow idle time, packet inter-arrival time and packet lengths). The main challenge to the implementation of such machine learning solutions so far has been the lack of ground truth, highly non-stationary data distribution and algorithms that underperform. There still lies an opportunity for developing highly accurate classifiers.

2.2.2 Spectrum use

AI can be determinant towards deciding on transmission schemes, access methods, carrier frequency and bandwidth, channel modeling and transmission power.

Concerning HFC networks, one of the most important MAC layer functions introduced by the DOCSIS 3.1 specification is profile management. For each cable modem (CM) a profile is set, meaning a set of subcarriers and modulation orders are configured, based on the channel signal-to-noise ratio (SNR). This aims to increase spectral efficiency. The profiles are defined by the CMTS, based on the relationships existing between SNR and modulation. There remains an opportunity for vendors and CSPs to work on AI algorithms and improve profile management.

Next generations of wireless networks depend heavily on channel state information (CSI), which is an estimate of the capacity of a communication link, used to optimize signal transmission. Current methods for the estimation of CSI entail considerable air interface resource overhead. It was discovered that exploiting linear correlations of CSI among, for example, co-located antennas, different time instances, and different frequency subcarriers can alleviate this problem. There exists potential in exploiting nonlinear CSI structures [19].

2.2.3 Proactive Network Maintenance

Degradation of the physical network can be assessed through proactive network maintenance (PNM) practices, which are developed from the analysis of data collected from the Customer Premise Equipment (CPE). This allows for the detection of several network impairments, such as impedance mismatches, internal noise, attenuation, signal ingress and egress, suckouts, among others.

PNM has been around for at least 10 years and has reached maturity. However, new capabilities continue to be developed. An example of this is the expansion of diagnostic measurements provided by DOCSIS 3.1 specification.

Many improvements in the diagnosis of the network are the result of applying machine learning techniques to the CPE data, in conjunction with other data sources. Additionally, the scope of the analysis can be expanded by including other data sources, to assess network reliability and predict the probability of impairment, which, as shown in [20] can be used to reprioritize tasks and so avoid the appearance of further damages.

2.2.4 Capacity planning

At Telecom Argentina, AI is determinant for capacity planning. In the case of the HFC network, we use time series to forecast traffic and resource utilization at total, per regional hub or local site, per service group and per subscriber levels. This leads to a data-driven design, and optimization of resources. Similarly, in the case of wireless networks, the same kind of longitudinal analysis is made at cell site, cell or antenna and subscriber level. Tuning network parameters enables us to cope with subscribers' requirements.

Regarding wireless networks, ML and other state-of-the-art techniques and technologies -such as digital twins- are key to prevent coverage holes. Information from sites and cities infrastructure can be combined and analyzed in order to improve coverage. The need for automated detection of holes is particularly relevant for the case of 5G millimeter-wave signals (mmWave), which are subject to random blockage.

Moreover, in 5G, the link capacity between users and base stations (BS) can be much higher compared to sub-6 GHz wireless systems. Meanwhile, due to the high cost of infrastructure upgrade, it would be difficult to drastically enhance the capacity of backhaul links between mmWave BS and the core network. As a result, the data rate provided by backhaul may not be enough to support all mmWave links; hence, the backhaul connection becomes the new bottleneck. BS-UE link is characterized by high data rate and unstable connection, while the backhaul link is characterized by relatively limited data rate and stable connection. To balance this mismatch and enhance the system performance, efficient backhaul resource allocation to each user is necessary. Such adaptive control cannot be implemented by traditional resource allocation schemes due to the varying system dynamics [21].

2.2.5 Beamforming

Beamforming technology aims to provide faster, stronger signals with longer range. In contrast to omnidirectional signal transmission, it consists of the broadcasting of signals in specific angles, in the form of beams. Provided that a higher number of active beams increases the system resource utilization, it becomes necessary to leverage system utilization through an optimal definition of the beams to be used. In this sense, AI could provide a solution.

User equipment (UE) measures the beam state information (BSI), which is based on measurements of beam reference signal (BRS), comprising of parameters such as beam index (BI) and beam reference signal received power (BRSRP). Given a set of potential beams, the best is the one with the highest signal strength a.k.a. RSRP. An algorithm could be trained to automatically find the best beam, considering the RSRP and/or the BI [17].

2.2.6 Massive MIMO

Multiple-input multiple-output (MIMO) is a wireless technology that leverages link capacity, by exploiting multipath propagation. Normally, transmitted signals bounce off walls, ceilings, and other objects, reaching the receiving antenna multiple times at different angles and slightly different times. MIMO technology uses multiple, smart transmitters and receivers with an added spatial dimension, increasing performance and range [22]. MIMO antennas usually have two transmitting and two receiving elements, which double the capacity of a basic antenna. Massive MIMO goes further, using multiple elements simultaneously.

The weights for antenna elements for a massive MIMO 5G cell site are critical for maximizing the beamforming effect. AI can be used to identify dynamic change and forecast the user distribution (based on historical data), dynamically optimize the weights of antenna elements, perform adaptive optimization of weights for specific use cases with unique user-distribution and improve the coverage in a multi-cell scenario considering the inter-site interference between multiple 5G massive MIMO cell sites [23].

Additionally, the implementation of massive MIMO implies the installation of many antennas in the base station antenna array, which requires many power amplifiers (PA). The primary problem with PAs is the existing trade-off between their capacity of nonlinearity tracking, predistortion, and impairment correction, and their electrical efficiency. Highest PA efficiency is achieved when constantly feeding the PA at the limit of its highest-power linear region. This is not a feasible solution for high peak-to-average power ratio (PAPR) signals and is not realistic for 5G BSs. Instead, signal-processing-based solutions (ML solutions) are used to provide a better cost-performance trade-off.

2.2.7 Encoding

In [24], the enhanced structure of a Deep Neural Network (DNN) based encoder and decoder was developed for orthogonal frequency-division multiplexing (OFDM) systems and a variation of the autoencoder structure was applied to build a codec for sparse code multiple access (SCMA) systems. There exists a theoretical possibility of using DNN-based communications systems, which can adapt their operation according to the surrounding environment, including, dynamic channel, mobility, power, and so on. The operation of DNN-based communications systems in practice, especially with respect to their system architectures and the implementation of hardware capable of real-time operation, has not yet been verified.

From the point of view of computational cost, building an encoder and a decoder with a DNN may be challenging because it requires a large number of arithmetic computations (i.e., multiplications and additions) for matrix operations in DNN. Therefore, in order to accelerate operations in the DNN-based encoder and decoder, there is a clear need for dedicated and pipelined digital circuit designs, in which the processing time can be reduced to support the high data rate.

2.2.8 NFV/SDN

The following notes are a synthesis of the survey in [25]. NFV allows customers to transfer the networking functions from vendor-specific and proprietary hardware appliances to software hosted on COTS plattforms [26].

Thinking of today, we have several kinds of VMs running on the same type of servers within the datacenter, and we still have routers and switches in different specific boxes. In the NFV context, a new generation of servers with network functionalities designed at the hardware level including also the application layer will be capable of running not only services inside VMs but also network functions such as firewall, routing and switching. This has a tremendous advantage in scaling.

SDN is a new paradigm that was designed to overcome the difficulty in developing and testing new solutions and protocols in production environments, where the underlying code running in business switches and routers are proprietary and closed [27].

The main feature of the SDN paradigm is the separation of the control and data planes. Centralization of the control plane in conjunction with the availability of open APIs, making easy the process of creating and deploying new network configuration and management. This is a useful and powerful abstraction that has as much implications as existing applications and protocols. Maybe one of the main challenges is not technical but strategic: how can two big companies which compete for the same market be ready to sit down, discuss and write code about common APIs?

Both concepts have not only new technical paradigms which are required by the community, they also have consequences at commercial strategy level of big companies. And both concepts are the first steps in order to apply AI models and algorithms in the broad sense inside the networks.

2.2.9 Wireless and Mobile Networks

Improved data rate, low latency and increased capacity for consistent QoS/QoE are the main drivers of 5G networks. The final standard will be published by the ITU in mid-2020, which is also referenced as International Mobile Telecommunications (IMT)-2020. The 5G Tech Forum was created with the participation of the major vendors such as Verizon, Cisco, Ericsson, Nokia and Apple in order to develop early 5G specifications. It is very important to follow what these people say, and how it can relate to the AI environment, thinking not only on applications but also in network algorithms running on the mobile device.

5G and the Cable Industry agree to provide features that support different types of vertical businesses such as IoT, Automotive, Health care, VR&AR, IPTV or Media & Entertainment, and for those applications the requirements are very different: while automotive industry will need very low latency, IPTV is also currently requiring more bandwidth, and so on.

The access networks are currently being transformed into a new convergent scenario: the converged access transport network (CATN), which blurs the boundaries between access and transport networks driven mainly by the need for bandwidth on existing networks. Radio Access Network (RAN) also must evolve supporting the coexistence of different radio access technologies such as LTE and Wi-Fi. The challenges about data processing are being covered by techniques such as Mobile Edge Computing (MEC).

The use of NFV and SDN technologies will play a significant role in 5G networks, since they allow the network programmability and the fast delivery of new services, enabling network slicing and MEC implementation and orchestration. So, imagine they are running, how we can monitor or operate this level of infrastructure? It is evident that the AI platforms have nowadays a list of requirements to be guaranteed in the next few years.

2.2.10 Network Slicing

Network Slicing refers to the partitioning of a certain physical infrastructure, composed of both network and computational resources, into multiple logical networks, called network slices. This approach has several advantages over the traditional physical networks, such as customization of logical networks according to service requirements, on-demand provisioning to scale resources up or down as conditions change and network resource isolation for improved security and reliability.

The strategy to keep in mind here with the AI approach is to be bendable: more flexible in terms of deployments and in case system fails, we can accept a partial degradation of the services, but we don't want a cut. So different ML algorithms to monitor, maintain the operators reported and make decisions in real time will be needed. Maybe this is a very important point in terms of security and audition of the AI approach. How can a system be audited if a wrong decision is taken by AI?

2.2.11 Mobile Edge Computing

MEC or Multi-access Edge computing has been a trend in mobile networks. The MEC architecture has been standardized by ETSI since 2016, providing IT and Cloud Computing capabilities within RAN. For this, a set of computer and storage resources are deployed at the edges of a mobile operator's network to assist the core data center in supporting computing and communication. It focuses on delivering the services closest to the user to meet certain critical application requirements that are not supported only by Cloud Computing, such as high bandwidth, low latency and jitter, context awareness, and mobility support.

There are studies that deal with Distributed NFV and multiple VIMs (Virtualized Infrastructure Managers) which goes in the complementary approach of the central control plane proposed by SDN. But there are interesting works with these D-NFV; for example, it includes a Virtual Network Life Cycle Manager (VNLM). This element implements a multi-objective resource scheduling algorithm that uses a genetic algorithm to provide near-optimal placement of VNFs over different data centers.

When we consider the use of multiple VIMs without distributed NFV, we usually have a scenario where there are two secondary VIMs, one to manage a data center for virtualization purposes and another to manage a transport network for end-to-end network services provisioning.

2.2.12 Optimizing Customer Experience on Video using Machine Learning

Telecom has its own IPTV and Streaming platform, called FLOW, in order to deliver the best entertainment service to its subscribers, to increase market penetration and to gain competitive advantage. This deployment is based on unmanaged (second screens) and also on managed devices (set top boxes).

To implement our ABR streaming services we need to use different types of video encodings. Encoding is a multi-layer matter of concern. While in DOCSIS we have encoding at the MAC layer, we do not have to forget that the success of video streaming (QoE) depends on the adequate trade-offs between encoding and available bandwidth. In the industry, objective and subjective models for measuring video quality are well known. Recently, objective models emerged using ML algorithms which are trained using databases with subjective evaluations, to combine a variety of classical metrics. Classical metrics are much simpler to implement and at a lower cost but, they produce worse results that do not always fit the human perception. On the contrary, the metrics based on AI produce results very close to the subjective opinion of the customers, but they are more complex to implement and provide development opportunities.

Within the objective metrics based on AI, there are two methods that produce similar results: Video Multimethod Assessment Fusion (VMAF) and Video Quality Model with Variable Frame Delay (VQM-VFD). VMAF is an open-source method proposed by Netflix in 2016 and VQM-VFD was standardized by ITU in 2003. Thus, it’s very important for us to develop VMAF as a tool for optimizing FLOW customer experience and to equalize video quality with the other existing video platforms.

3 Technical and Theoretical Framework

In this section we comment on the roadmap of AI for Network Operations in Telecom Argentina. We mention some ML algorithms, as well as some others that are non-ML. Both types shape an ecosystem of techniques and technologies that coexist to make the development of data-driven solutions possible. Far from thinking these methods have to be part of a standard, we mention the algorithms for which we have found interesting applications.

3.1 Machine Learning

The following are well-known algorithms, so rather than focusing on the methodology itself, we enphasize on the practical problems, how they are posed in terms of variables, how these methods provide a solution and the practical implications. Some of the applications are thoroughly documented in publications that we, as members of the STEM team, have submitted to conferences and industry events.

- Artificial Neural Networks (ANN)
- Time Series models
- A/B Testing
- Clustering
- K-Nearest Neighbors (KNN)
- Context Aware Recommendation
- Natural Language Processing (NLP)
- Deep Learning techniques

3.1.1 Artificial Neural Networks

Maybe one of the most extended approaches in the application of Artificial Neural Networks (ANN) is the convolutional architecture (CNN), which is originally designed for image recognition and extended for natural language processing, speech processing, and computer vision. The main concept behind these structures consists of the application of a series of transformations to a set of independent variables, in order to predict a response or objective variable (Figure 5). As part of Telecom Argentina STEM team, we have used them to classify the outside plant investments of the optical nodes in the HFC access [28]. Additionally, Marketing teams applied them to predict churn.

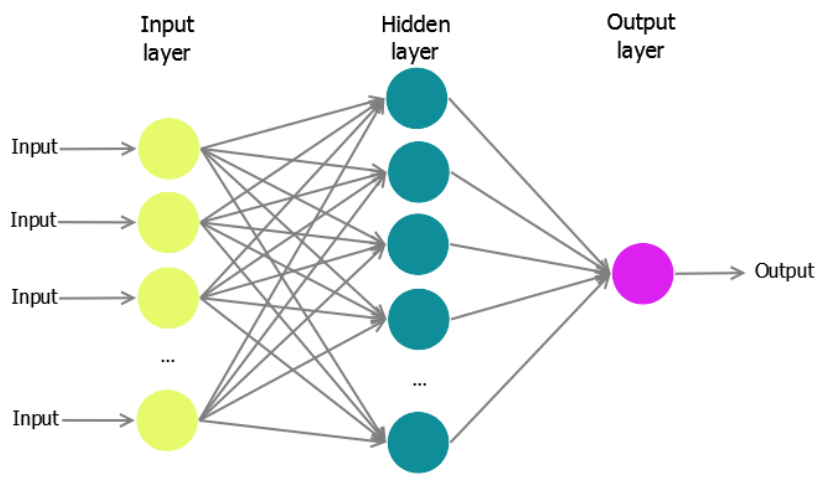

Figure 5 - Artificial Neural Network conceptual diagram

Other architectures were explored for a variety of purposes. For anomaly detection and forecasting of time series, such as CPU load or traffic growth respectively, the Long-Short Term Memory (LSTM) architecture has been tested effective [29].

3.1.2 Time Series models

Time series are often used to forecast the global traffic growth in the industry. Although it seems there is a consensus in the industry about the use of linear regression as the forecast method even when the hypothesis of independency between observations is not met, advanced forecast methods such as ARIMA, exponential smoothing or Kalman filter are, in our experience, more accurate and they have been proved to be useful to explain auto-dependence, stochastic trends and seasonality [30].

3.1.3 A/B Testing

A/B Testing is often a rigorous way to determine the results of a change in productive environments. It is an empirical approach in which an experiment is conducted considering two groups of subjects: one is the control and the other, the test. By definition, the groups are mutually exclusive. A certain change or alteration is introduced to the test subjects, so its consequences are measured within the test group and compared to control (which was not subject to changes). The results are driven by the actual data, ending with subjectivities. An application example is testing two different posters for a given TV channel, which are evaluated to determine the the best promoter. A/B Testing will be one of the fundamental keys in the operations toolbox of the future networks, because it allows to test network changes in production environment at lower costs than current migrations.

3.1.4 Clustering

Clustering techniques aim to assign analysis units to classes, according to the appearance of similar characteristics among them. In other words, for a dataset where we have observations and features, cluster analysis is used to find groups such that the observations within the same group are close in the given features space. There exist various definitions of distance, being applied most frequently the Euclidean and Manhattan distances or other correlation-based statistics such as Pearson, Kendall and Spearman correlation coefficients.

There are three main types of algorithms: hierarchical, density-based and non-hierarchical. Hierarchical clustering are useful to find groups and sub-groups within them. Density-based methods do not assume that every observation belongs in a cluster, instead they classify some of the cases as clusters and some others, as 'noise'. Non-hierarchical methods are particularly useful when a priori we can assume that clusters have similar size. In our experience, the latter have been useful to find subscribers with common trends in geographical location and service tier acquired. Another application we found for the same type of methods is to classify cable modem signals, according to the appearance of impairments. This is helping in the development of further capacities for PNM [31].

3.1.5 K-Nearest Neighbors

K-Nearest Neighbors (KNN) is a supervised algorithm that can be used for classification as well as for regression problems. Provided a certain objective variable and set of p features, which determine a p-dimensional feature space, we know the location of the points from a training sample and their response values. When a new observation comes along, and we want to predict its response value, the strategy is to look at the response of the K closest observations in the p-dimensional feature space, and use them to infer on the new one. The method presents the advantage of being non-parametric and robust. In counterpart, it is computationally expensive to search for the nearest neighbors for all observations in the sample. At Telecom Argentina, this is used to recommend VoD contents to subscribers, based on what other similar users who have watched in the past.

3.1.6 Natural Language Processing

The processing of natural language, also known as NLP, is a branch of computer science associated to the traditional AI definition, which aims to give computers the capacity to understand the meaning of human language. There are efforts directed towards analyzing language according to formal definitions and relationships, executing grammatical, semantical and syntactical analysis. On the other hand, the statistical approach, which consists on extracting words statistics, has been much more successful. One of the reasons for this is when an analysis tries to go deeper by considering relationships between entities and implications to the real world, it loses potential to generalize to other application areas.

Sentiment analysis is one of the most common aplications of NLP based on social networks comments in order to evaluate the social experience of the services. Apart from that case, we made some work analyzing the descriptions and comments of the people who monitor the network (like the NOC), to check if the failures or disruptions of the network are being manage uniformly and how they appeared repeatedly along time.

3.1.7 Deep Learning techniques

These notes are based on the article [32]. Deep learning uses multiple layers to represent the abstractions of data to build computational models. Deep learning, which has its roots from conventional neural networks, significantly outperforms its predecessors. It utilizes graph technologies with transformations among neurons to develop many-layered learning models.

Feature engineering focuses on building features from raw data and is often very domain specific and requires significant human effort. Some very well-known features proposed to compare are Histogram of Oriented Gradients (HOG), Scale Invariant Feature Transform (SIFT), and Bag of Words (BoW).

Deep learning algorithms perform feature extraction in an automated way, which allows researchers to extract discriminative features with minimal domain knowledge and human effort.

Maybe the most prominent current features about Deep Learning techniques are the possibilities given by the capacity of parallel and distributed computing techniques which makes viable lot of implementations.

Threshold logic is the combination of algorithms, mathematics and architectures which allows to emulate the process of thinking of humans, but not to learn in the human sense. The perceptron is the first device within the context of cognition systems, and we all know about how lots of perceptron interconnected conforms a neural network.

The backpropagation algorithm uses the errors in training deep learning models and is one of the first milestones in the creation of neural network architectures. Then the Recurrent Neural Networks (RNN), Recursive Neural Networks (RvNN) such as the Long-Short Term Memory RvNN (LSTM), Deep Neural Networks (DNN), Deep Belief Networks (DBF), Restricted Boltzmann Machines (RBMs) and others are some well-studied architectures.

In each problem we face with Artificial Neural Networks (ANN), we deal with architectures, hyperparameters, optimization algorithms, loss functions, and most important, the nature of the data domain. The mature of this technology is in direct relation with the performance achieved using the pipeline process of the GPUs for processing data vectors instead of finishing in pixels, and since that milestone, several frameworks such as TensorFlow, Torch, Theano, MXNet and others were developed and currently have a big community of developers.

3.1.7.1 Deep Learning in Distributed Systems

There are two main approaches to train models in a distributed system: data parallelism and model parallelism. In the former the model is replicated to all the computational nodes and each model is trained with the assigned subset of data; in the latter all the data is processed with one model where each node is responsible for the partial estimation of the parameters in the model.

With parameter averaging we have for example N slave nodes and one master node, and at time t the weight on the master node is Wt, then:

$$W_{t+1} = \frac{1}{N}\sum_{i=1}^{N} W_{t+1,i}$$

is the weight at time t+1. Very often, the objectives of a ML algorithm are optimized using the update:

$$w \leftarrow w - \alpha \sum_{i=1}^{n} g(w; x_i, y_i)$$

where w is a vector of dimension d and the data has a length of n, x_i and y_i are the dimensions or hyperparameters of the model and alpha is the weight.

These simple notations are placed here to mention performance trade-offs between memory and computing cost and new compromises driven by the applications.

If we have MEC the ML models must be solved within the capacity of the edge, so only local data parallelism is admitted. But if we have a centralized computation cluster, then we can apply global models and data parallelism for the whole network, but at the price of latency for applications running far away from the datacenter.

So, the implementation of AI solutions on future networks will have to address this compromise too. MEC is a powerful framework to focus on the models and applications very near the user. This will also require the need for monitors and fail recoveries schemes when, for example, some application is compromising the capacity of the local cluster.

3.2 Non-Machine Learning Tools

As we proceeded with traditional ML algorithms, next we describe some other tools that are not ML, since their implementation does not require tuning parameters, optimization of functions or training models. However, they complement the ML implementations, in some cases they provide contextual information and in some others, they are useful to execute some calculations with reasonable timing.

3.2.1 Locality-Sensitive Hashing

Locality-sensitive hashing (LSH) facilitates the identification of similar observations in a high volume of data. The general idea is to hash data points into buckets so that data points near each other are located in the same buckets with high probability, while data points far from each other are likely to be in different buckets. While the traditional process of looking for pairs of equals has a quadratic computational cost, $O(n^2)$, this alternative looks for pairs of similar items without the cost of examining each possible pair, henche a much lower cost. It is particularly useful to detect near-duplicate documents, webpages and other types of files, for large-scale image search and audio/video fingerprinting (A/V fingerprinting) [33]. Fingerprinting enables the identification of characteristics from multimedia, so we expect to use this algorithm to extract features from contents in Telecom Argentina's Flow platform.

3.2.2 Collaborative filters

Collaborative filtering is used for content recommendation. It is based on previous subscriber behavior, and it evaluates similarities between subscribers in terms of the viewed content. An index score for each content is defined with the measurements of each user and then is multiplied by the similarities distance between users. The similarity analysis is usually combined with the results of a clustering algorithm to improve recommendations [34]. In the case of Telecom Argentina, collaborative filtering is used as a complement to KNN clusters to manage recommendations of VoD content.

3.2.3 Process Mining

When we execute an analysis, we intend to answer the questions what happened, why it happened, what is likely to happen next and what is the best that could happen next. The analysis of processes allows us to understand if a certain process model that is assumed is actually being executed, as well as discover new process models underlaying the data. Sequences and relationships between possible events are represented with Petri nets, where nodes represent the events, directed arcs indicate the pre-conditions and postconditions, and they can be enriched with information about time between events or frequency of appearance [35].

In a digital ecosystem, where every device produces event data (in the form of data logs), we expect that process mining (PM) will provide an alternative to understand subscriber behavior and compare the expected usage patterns to actual ones.

3.2.4 Digital Twins

Digital twins are realistic representations of real-world entities, which consist of a combination of mathematical and computational methods, and software services that provide real time synchronization with the real object or process they represent. Their purpose is to judge, analyze, predict and optimize the real entity. They are usually supported by 3D, video, augmented reality and/or virtual reality representations.

The twins are used in manufacturing to reduce errors, reduce planning time, plan for changes, hence reduce cost of changes, and increase the planning maturity [36]. Another field of application is smart cities, as it is the case of the digital twin created for the city of Atlanta [37], to plan for resource allocation, provide security, maximize services, facilitate human activities and prevent disruption while continuously adapting. We expect that the same approach can contribute to improve network capacity planning, engineering and architecture.

3.3 The Knowledge Plane

The research community has considered in the past the application of AI techniques to control and operate networks. For example, in 2003 David Clark et. al propose the knowledge plane (KP) as a *pervasive system within the network that builds and maintains high level models of what the network is supposed to do, in order to provide services and advice to other elements of the network. The knowledge plane is novel in its reliance on the tools of AI and cognitive systems* [38].

The knowledge plane (Figure 6) paradigm proposes the evolution to a cognitive network, where the devices learn, decide, and act to achieve end-to-end goals. This emerging paradigm is clarifying a set of new cognitive-based protocols and algorithms that optimize network's performance.

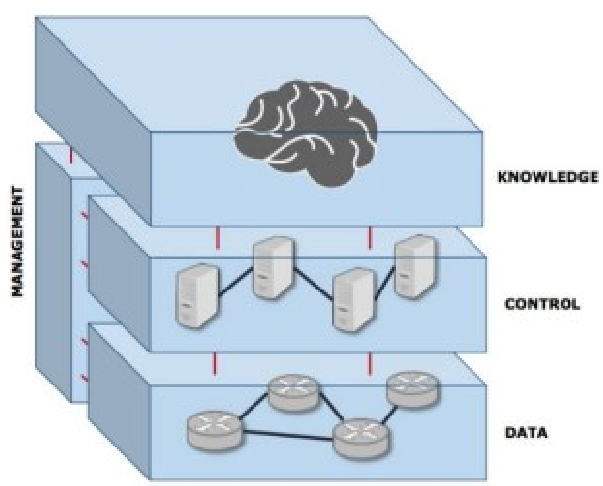

Figure 6 - The four planes in network architecture.
Source: TM Forum.

Several different KP based approaches have been proposed [39]. But it is not until the development of NFV and SDN that such proposal once again takes hold in communities such as the TM Forum, IETF and the Industry.

In [40] progress is made in the definition of a new paradigm based on this plane. This is knowledge-defined network (KDN) operates by means of a control loop to provide automation, recommendation, optimization, validation and estimation. The KDN paradigm is also taken by the TM Forum as a proposal to specify future architectures [41].

An example of the plane of knowledge in an SDN is presented in Figure 7 [42].

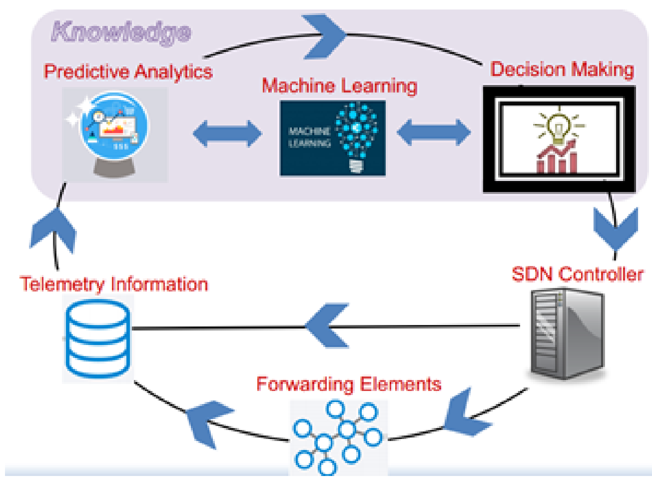

Figure 7 - Plane of Knowledge in a SDN

Finally, Figure 8 shows an architecture with more detail presented in APNOMS [43].

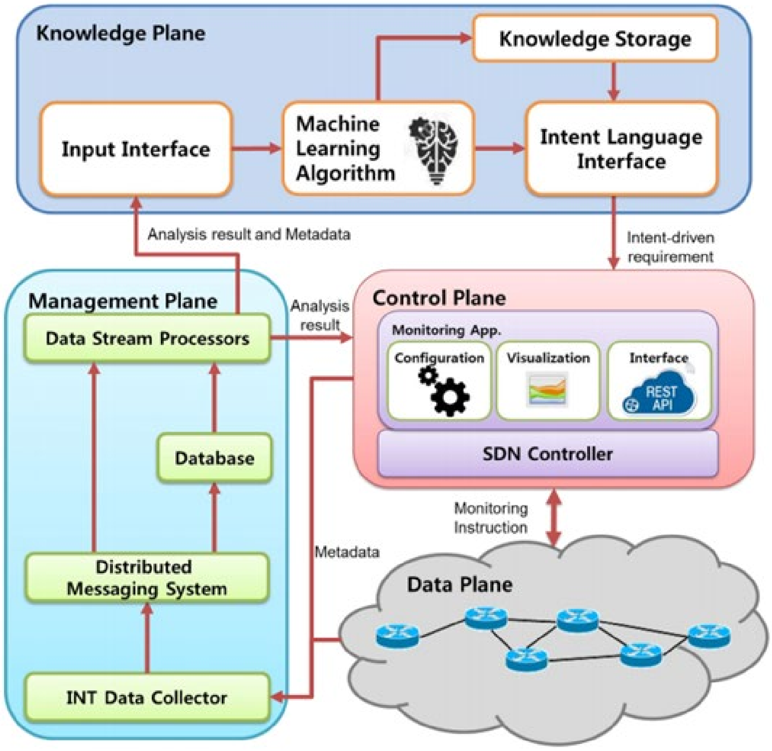

Figure 8 - Detailed Plane of Knowledge in a SDN

4 Strategy

More than a decade has passed since the emergence of a paradigm of autonomous computing in the world of telecommunications. Back then, there was a gap between that paradigm and the capacities of the networks. However, a path has been taken in recent years with the adoption of cloud computing, NFV and SDN.

These technological advances have made available a more agile infrastructure, computing capacity and storage as resources more abundant than ever. Motivated by this evolution, together with the ever-growing need to improve the management and administration of networks and services, we present this first approach to the plane of KP.

We define a reference architecture (Figure 9) with the purpose of automating the services from end to end, with a holistic view of the network and towards a CATN. It is based on RFC 8309, MEF-55, eTOM model (TM Forum) and other initiatives of other CSPs and vendors.

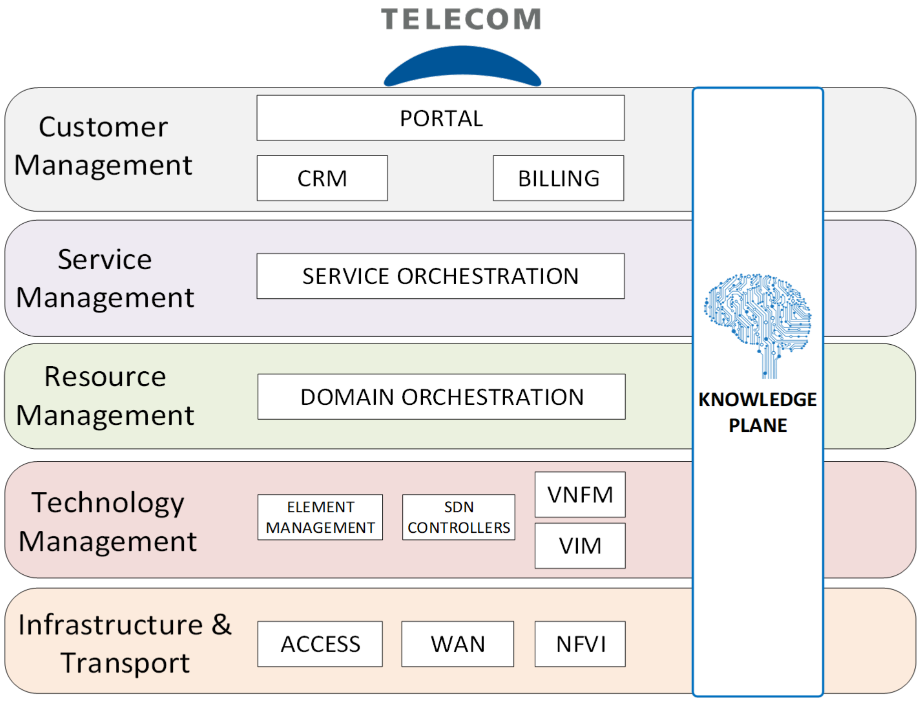

Figure 9 - Telecom Argentina knowledge-defined network

We detail some of the main topics around AI and networks. The first steps Telecom Argentina is giving towards the deployment of AI for networks and services are:

- Operations use cases of Deep Learning

The main benefits of the deployment of AI in the context of networks will be achieved by the operations teams. This is often the main starter point to think about a deployment of ML. One of the well-known use cases is anomaly detection.

- NFV/SDN orchestration layer

We consider SDN orchestration layer as the basis for all the AI running applications. But this does not mean we will wait for it before starting with AI.

- Extension of AI applications through integration platforms

Today we are used to having several OSS/BSS systems that are not always interconnected within a common platform. In this sense, integration platforms are key technologies which enable the combination of different data domains and kinds of algorithms through common APIs. Some relevant work about integration platforms with native parallelism has been made by the people of The Apache Software Foundation, particularly with Spark.

- Highly specialized human resources

The structure of the organization is another key concept which cannot be excluded by the AI framework. The knowledge layer is a concept in the SDN paradigm with relevant meaning in the ability of the specialized teams to deeply understand about topics of networks, machine learning and business.

- Distributed Data Lakes and centralized Data Ocean

There are a lot of data domains available in the current deployed networks, but this does not mean all the data domains have to be in a centralized datacenter. There are centralized and distributed approaches for the administration and mining of the data, and a lot of effort to build a data ocean. But as the capillarity of the networks is still growing and merging with the mobile access, the MEC could be evolve as a network of distributed data lakes.

- TM Forum framework for application and analytics layers

Hybrid layer models considering the enterprise organizations have been designed consider mainly TMForum for the application and analytics layers and ETSI/IETF for the Physical and MAC layers.

- Open source and microservices-based platforms

Open source is useful at the development and test stages, and sometimes with the adequate support is the right tool for production environments. Vendors are taking that direction to make better products and to guarantee a dynamic evolution and technical support.

- Data Governance

Data governance is still developing the interactions and processes within the organization. The main capabilities are to manage data quality, security and compliance, storage, presentation and distribution of the data within and outside the enterprise. It also has to meet compliance requirements dictated by the organization policies and external regulatory entities [44].

Conclusions

The Applied Artificial Intelligence STEM Group will focus on the application of AI to decision-making process and auto-remediation to help Telecom Argentina's network keep pace with the growth in network size, traffic volume and service complexity, as well as define new approaches to network operations and customer assurance to support the accelerated deployment of new over-the-top services and collaborate in the digital transformation process, in which Telecom Argentina goes from being an MSO & MNO to a Digital Service Provider.

This work is the first step in the Telecom Argentina's roadmap for AI applied to networks and services. It is a long, challenging journey, however, we are confident that by taking one step at the time, little but concrete, we will accomplish our goals.

The first stage we are working on is to define this framework regarding to the deployments of AI systems in the whole network. More than a year ago we understood that there will be no future networks without AI and we decided to make our first proof of concepts based on our network and services data. We already know we have to continue the direction of this decision but this does not mean we have to stop testing models and algorithms in our labs. Some methodologies like Agile or DevOps are very useful in order to find fast results.

The trainings are some of the most important skills the knowledge layer must develop. Customized trainings based on local data and exposed by local professionals are a good practice.

Working groups such as CableLabs/SCTE, TM Forum and Telecom Infra Project are collaborative environments the industry may maintain, and the company will continue to participate.

Regarding the vendor relations, in this stage we are identifying how they deploy the knowledge layer and how is promoted. We already define a convergent scenario for the merging of the access networks called Converged Access Transport Network (CATN), so the vendor's offered knowledge layer has to match with our requirements. This is a proposal to invite you to follow our path.

Abbreviations

ABR	adaptive bit rate
AI	Artificial Intelligence
ANN	Artificial Neural Network
AR	augmented reality
ARIMA	autoregressive integrated moving average
Avg	average
B2B	business to business
B2C	business to consumer
BS	base station
BW	bandwidth
CAGR	Compound Annual Growth Rate
CM	Cable Modem
CMTS	Cable Modem Termination System
CVA	Cablevisión S.A.
DNN	Deep Neural Network
DOCSIS	Data Over Cable Service Interface Specification
DSP	Digital Service Provider
EDA	Exploratory Data Analysis
GHz	Giga Hertz
HFC	Hybrid Fiber Coaxial
HHP	household passed
IoT	Internet of things
IPTV	Internet Protocol Television
ISBE	International Society of Broadband Experts
Kbps	kilobits per second
Km	kilometers
KPI	key performance indicator
LTE	long term evolution
Mbps	megabits per second
MEC	multi-access edge computing
MIMO	multiple input, multiple output
ML	Machine Learning
mMIMO	Massive MIMO
MNO	Mobile Network Operator
MSO	Multiple Service Operator
NFV	network functions virtualization
NR	new radio

OPS	operations per second
O-RAN	Open Radio Access Network
OSS/BSS	Operation Support System/Business Support System
PCA	Principal Components Analysis
PNM	Proactive Network Maintenance
QAM	Quadrature Amplitude Modulation
QoE	quality of experience
QoS	quality of service
RAN	radio access network
ReLU	Rectified Linear Unit
SCTE	Society of Cable Telecommunications Engineers
SD	standard deviation
SDN	software-defined network
SG	service group
SoC	system on a chip
SON	self-organizing network
STEM	science, technology, engineering and mathematics
Subs	subscribers
TCO	total cost of ownership
UE	user equipment
URLL	ultra-reliable low-latency
URLLC	ultra-reliable low-latency communications
VR	virtual reality

Bibliography & References

[1] «Digital Service Providers (DSP),» National Cyber Security Centre (NCSC), 2019. [En línea]. Available: https://www.ncsc.gov.ie/dsp/.

[2] J. Kyriakakis, «Defining the Digital Service Provider,» LightReading, 29 September 2014. [En línea]. Available: https://www.lightreading.com/business-employment/business-transformation/defining-the-digital-service-provider/a/d-id/711114.

[3] J. Bryson y A. Winfield, «Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems,» *IEEE Computer Society,* vol. 4, nº 2, pp. 10-13, 2018.

[4] A. M. Turing, «Computing Machinery and Intelligence,» *Mind,* vol. 59, nº 236, pp. 433-460, 1950.

[5] J. McCarthy, «What Is Artificial Intelligence? - Stanford University,» 12 November 2007. [En línea]. Available: http://www-formal.stanford.edu/jmc/whatisai/node1.html.

[6] S. Russell y P. Norvig, Artificial Intelligence: A Modern Approach, 3rd Ed., Prentice Hall Series in Artificial Intelligence, 2010.

[7] A. L. Samuel, «Some Studies in Machine Learning Using the Game of Checkers,» *IBM Journal of Research and Development,* vol. 3, nº 3, pp. 210-229, 1959.

[8] B. Marr, «How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read,» Forbes, 21 May 2018. [En línea]. Available: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#2aff0a3960ba.

[9] K. Schwab, The Fourth Industrial Revolution, Penguin Random House, 2017.

[10] C. Righetti, *Personal notes on TM Forum Digital Transformation,* Nice, France, 2018.

[11] K. Sundaresan, N. Metts, G. White y A. Cabellos-Aparicio, «Applications of Machine Learning in Cable Access Networks,» de *SPRING TECHNICAL FORUM, CableLabs SCTE NCTA 2016 Spring Technical Forum Proceedings*.

[12] G. White y K. Sundaresan, «DOCSIS 3.1 Profile Management Application and Algorithms,» de *SCTE NCTA 2016 Spring Technical Forum Proceedings*.

[13] H. Hagras, «Toward Human-Understandable, Explainable AI,» *IEEE Computer,* vol. 51, nº 9, pp. 28-36, 2018.

[14] B. S. I. (BSI), British Standard BS 8611:2016, Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems., London, UK, 2016.

[15] U. Gasser y V. A. F. Almeida, «A Layered Model For AI Governance,» *IEEE Internet Computing,* vol. 21, pp. 58-62, 2017.

[16] T. Forum, «IG1184 Service Management Standards for AI R18.5.1,» 2019.

[17] M. Yao, M. Sohul, V. Marojevic y J. H. Reed, «Artificial Intelligence Defined 5G Radio Access Networks,» *IEEE Communications,* vol. 57, nº 3, pp. 14-21, 2019.

[18] T. T. T. Nguyen y 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.

[19] Z. Jiang, S. Chen, A. F. Molisch, R. Vannithamby, S. Zhou y Z. Niu, «Exploiting Wireless Channel State Information Structures Beyond Linear Correlations: A Deep Learning Approach,» *IEEE Communications Magazine,* vol. 57, nº 3, pp. 28-34, 2019.

[20] L. Wolcott, M. O'Dell, P. Kuykendall, V. Gopal, J. Woodrich y N. Pinckernell, «A PNM System Using Artificial Intelligence, HFC Network Impairment, Atmospheric and Weather Data to Predict HFC Network Degradation and Avert Customer Impact,» de *SCTE/ISBE 2018 Fall Technical Forum*, Atlanta, GA, 2018.

[21] M. Feng y S. Mao, «Dealing With Limited Backhaul Capacity in Milimeter-Wave Systems: A Deep Reinforcement Learning Approach,» *IEEE Communications Magazine,* vol. 57, nº 3, pp. 50-55, 2019.

[22] Intel, «Learn about Multiple-Input Multiple-Output,» 25 March 2019. [En línea]. Available: https://www.intel.com/content/www/us/en/support/articles/000005714/network-and-i-o/wireless-networking.html.

[23] O. Dharmadhikari, «Leveraging Machine Learning and Artificial Intelligence for 5G,» Informed blog by CableLabs, 18 June 2019. [En línea]. Available: https://www.cablelabs.com/leveraging-machine-learning-and-artificial-intelligence-for-5g.

[24] M. Kim, W. Lee, J. Yoon y O. Jo, «Toward the Realization of Encoder and Decoder Using Deep Neural Networks,» *IEEE Communications Magazine,* vol. 57, nº 5, pp. 57-63, 2019.

[25] M. S. Bonfim, K. L. Dias y S. F. L. Fernandes, «Integrated NFV/SDN Architectures: A Systematic Literature Review,» *ACM Computing Surveys,* vol. 51, nº 6, 2019.

[26] European Telecommunications Standards Institute (ETSI), «Network Functions Virtualisation - An Introduction, Benefits, Enablers, Challenges and Call For Action.,» de *SDN and OpenFlow World Congress*, Darmstadt, Germany, 2012.

[27] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker y J. Turner, «OpenFlow: Enabling Innovation in Campus Networks,» *ACM SIGCOMM Computer Communication Review,* vol. 38, nº 2, pp. 69-74, 2008.

[28] C. Righetti, E. Gibellini, F. De Arca, C. G. Carreño Romano, M. Fiorenzo, G. Carro y F. R. Ochoa, «Network Capacity and Machine Learning,» de *SCTE•ISBE Cable-Tec Expo 2017*, Denver, CO, 2017.

[29] C. G. Carreño Romano y N. Clivio, «Sizing Techniques Applied to Network Capacity Planning,» de *IEEE Biennial Congress of Argentina (ARGENCON)*, Tucumán, Argentina, 2018.

[30] R. Chrobok, «Theory and Application of Advanced Traffic Forecast Methods,» Duisburg-Essen, Germany, 2005.

[31] Gibellini, E.; Righetti, C., «Unsupervised Learning For Detection Of Leakage From The HFC Network,» de *ITU Kaleidoscope 2018: Machine Learning for a 5G Future*, Santa Fe, Argentina, 2018.

[32] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. Presa Reyes, M. L. C. S. C. Shyu y S. S. Iyengar, «A Survey on Deep Learning: Algorithms, Techniques and Applications,» *ACM Computing Surveys,* vol. 51, nº 5, p. Article 92, 2018.

[33] S. Gupta, «Locality Sensitive Hashing - Towards Data Science,» Towards Data Science, 29 June 2018. [En línea]. Available: https://towardsdatascience.com/understanding-locality-sensitive-hashing-49f6d1f6134.

[34] J. Leskovec, A. Rajaraman y J. D. Ullman, Mining of Massive Datasets, Cambridge University Press, 2010.

[35] W. Van der Aalst, Process Mining, Berlin, Germany: Springer-Verlag Berlin Heidelberg, 2016.

[36] F. Biesinger, D. Meike, B. Kraß y M. Weyrich, «A Case Study for a Digital Twin of Body-in-White Production Systems,» de *IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, Torino, Italy, 2018.

[37] N. Mohammadi y J. E. Taylor, «Smart City Digital Twins,» de *IEEE Symposium Series on Computational Intelligence (SSCI)*, Atlanta,GA, 2017.

[38] D. D. Clark, C. Partridge, J. C. Ramming y J. T. Wroclawski, «A Knowledge Plane For The Internet,» de *Conference on Applications, Technologies, Architectures and Protocols for Computer Communications (SIGCOMM)*, New York, NY, 2003.

[39] K. R. Sollins, «An Architecture for Network Management,» de *Workshop on Re-Architecting the Internet (ReArch)*, New York, NY, 2009.

[40] 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 y C. Cassar, «Knowledge-Defined Networking,» *SIGCOMM Computer Communications,* vol. 47, nº 3, pp. 2-10, 2017.

[41] B. Levy y B. Graham, «TM Forum Future Architecture Strategy,» 2017. [En línea]. Available: https://www.tmforum.org/wp-content/uploads/2017/09/TM-FORUM-FUTURE-ARCHITECTURE-STRATEGY-v5_final.pdf.

[42] Z. Zhu, «Knowledge-Defined Network Orchestration in a Hybrid Optical/Electrical Datacenter Network,» de *Conference on Optical Network Design and Modeling*, Dublin, Ireland, 2018.

[43] J. Hyun y J. Won-Ki Hong, «Knowledge-Defined Networking Using In-Band Network Telemetry,» de *Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Seoul, Korea, 2017.

[44] TM Forum, «Frameworks Technical Report, Data Governance Functions and Implementations (TR261 Release 16.0.1),» 2016.

