electronics

Article

Implanting Intelligence in 5G Mobile Networks—A Practical Approach

Sumbal Malik [1,2], Manzoor Ahmed Khan [1,2], Aadam [1], Hesham El-Sayed [1,2,*], Jalal Khan [1] and Obaid Ullah [1]

[1] College of Information Technology, United Arab Emirates University, Abu Dhabi 15551, United Arab Emirates
[2] Emirates Center for Mobility Research (ECMR), United Arab Emirates University, Abu Dhabi 15551, United Arab Emirates
* Correspondence: helsayed@uaeu.ac.ae

Abstract: With the advancement in various technological fronts, we are expecting the design goals of smart cities to be realized earlier than expected. Undoubtedly, communication networks play the crucial role of backbone to all the verticals of smart cities, which is why we are surrounded by terminologies such as the Internet of Things, the Internet of Vehicles, the Internet of Medical Things, etc. In this paper, we focus on implanting intelligence in 5G and beyond mobile networks. In this connection, we design and develop a novel data-driven predictive model which may serve as an intelligent slicing framework for different verticals of smart cities. The proposed model is trained on different machine learning algorithms to predict the optimal network slice for a requested service resultantly assisting in allocating enough resources to the slice based on the traffic prediction.

Keywords: 5G; network slicing; machine learning; random forest; neural network

Citation: Malik, S.; Khan, M.A.; Aadam; El-Sayed, H.; Khan, J.; Ullah, O. Implanting Intelligence in 5G Mobile Networks—A Practical Approach. *Electronics* **2022**, *11*, 3933. https://doi.org/10.3390/electronics11233933

Academic Editor: Rameez Asif

Received: 12 October 2022
Accepted: 4 November 2022
Published: 28 November 2022

1. Introduction

As the world is transiting to digital cities, the issues with today's networks become a never-ending effort and are even getting more intense as networks evolve to cater for the applications of low latency, high bandwidth, massive connectivity, support for high mobility, etc. The complex and extremely high dynamic nature of the envisioned network configuration asks for reduced human intervention. Programmable networks were one of the initial attempts to remove humans from the equation. However, implanting intelligence in the networks for complex use cases of different smart city verticals need more than just achieving programmable networks. The features that 5G and beyond pledges provide are the key ingredients. Machine Learning and artificial intelligence (AI) based approaches further add to introducing human context and decision-making in the networks.

The telecom industry has been completely revolutionized by several developing technologies that enable new business models and provide customers with new experiences. The emergence of programmable systems such as Software Defined Networks (SDN) and Network Function Virtualization (NFV) have evolved and benefited the networks. Among the crucial services that 5G networks would encompass are autonomous driving, enterprise business models, virtual reality solutions, industrial automation, remote monitoring, smart health, smart cities, and many more. Network slicing (NS) is viewed as a major enabling technology for 5G by the Third Generation Partnership Project (3GPP). It would enable operators to effectively run several network instances over a single infrastructure to serve various applications, use cases, and business services while providing the highest possible quality of service (QoS).

NS, defined by Next Generation Mobile Network Alliance (NGMN) [1] is an important component of 5G networks enabling the coexistence of multiple isolated and independent virtual networks (slices) running on the same physical infrastructure. Each network slice is an independent self-contained virtualized end-to-end network with its virtual resources, topology, traffic flow, and provisioning rules which allows the network operators to run different deployments based on different architectures in parallel. Network slicing provides multifold advantages such as (i) provides multi-tenancy therefore, multiple virtual network operators can share the same physical infrastructure which reduces the capital expenses of deploying and operating new networks; (ii) can create customized slices with varying QoS (e.g., data speed, reliability, latency, delay, throughput) requirements for each service meeting the service level agreements of the services; (iii) supports on-demand cost-effective creation, modification and annul of slices increasing the flexibility and adaptability of the network management [2]. However, some of the major challenges of NS are (i) dynamically adapting application service quality requirements and (ii) unpredictable resource and service quality demands. Furthermore, it is also worth highlighting that NS implementation deals with routing issues in different settings. There are known challenges to routing [3], and the research community have proposed various solution approaches to readers such as interfere-aware routing [4], measurement-based routing, channel-diversity-based routing, priority-based multi-path routing, and experience-driven model-free deep reinforcement learning-based approaches [5], etc. On similar lines, the shortest path and route problem in the network is one of the major challenges in the transportation and communication sectors. In this connection, Karimnejad et al. [6] proposed a genetic, particle swarm optimization algorithm to find the shortest path in a network with mixed fuzzy arc weights due to the complexity of the addition of various fuzzy numbers for larger problems. Capri et al. [7] also proposed a fuzzy-based Ant Colony Optimization algorithm to solve the shortest path problems with different types of fuzzy weights. Sori et al. [8] implemented the elite artificial bees' colony algorithm to resolve the route planning challenges in robotics. In another study, Sori et al., [9] also implemented an algorithm to solve the challenge of finding a path with the lowest cost where the traversal time of the path does not exceed a predetermined time-bound. For ready reference, the readers are also recommended to look into the following research [10–13]. The enabling technologies to deploy network slicing are SDN, NFV, and cloud computing discussed in Section 2.

This paper is organized into six sections. Section 2 provides a tutorial to assist readers with an overview of 5G enabling technologies and associated concepts. Section 3 highlights the importance of integrating the various enabling technologies and concepts that are important to realize network slicing. Section 4 discusses the potential ways towards 6G. Section 5 discusses the crucial challenges of network slicing and develops a predictive model to predict the best network slice for a requested service. Finally, Section 6 elucidates the conclusion of the study.

2. Tutorial & Background

This section focuses on equipping the readers with the necessary background information on 5G enabling technologies, which is needed to comprehend the contents of this paper.

2.1. ETSI NFV MANO Reference Architecture

The ETSI introduces the NFV MANO architecture [14] a fundamental base solution to build the softwarized 5G networks. The 3GPP leverages this mechanism to create the network slices enabling the dynamic deployment of network functions as virtual network functions (VNFs) completely decoupled from the underlying hardware running these functions. The VNF life cycle management together with dynamic resource allocation to the network slices implemented as a set of interconnected VNFs are performed and managed by ETSI NFV Management and Orchestration (MANO). MANO is mainly responsible to manage and orchestrate the network services, VNFs, and other resources such as networking, computation, and storage.

The ETSI NFV MANO architecture is comprised of three main building blocks:

- **Management and Orchestrator (MANO):** responsible for managing and allocating resources to NFV architecture such as creation and deletion of slices, slice selection and resource allocation. The MANO architecture has three main components: *(i) Virtualized Infrastructure Manager (VIM):* controls the physical and virtual infrastructures. Furthermore, it manages the interaction of a VNF with computing, storage and network resources; *(ii) VNF Manager:* configure and manages the VNFs and takes care of the life cycle of the VNFs. A VNF manager can be deployed for each VNF or a single VNF manager can serve multiple VNFs; *(iii) Orchestrator:* manages the VIM and the VNF manager.

- **Virtualized Network Functions (VNFs):** VNF is the software implementation of network functions deployed on one or more virtual machines running over the common NFVI. A set of VNFs executed in a specific order is referred to as Service Function Chain (SFC). The MANO can communicate with VNFs via the Ve-Vnfm connection point.
- **NFV Infrastructure (NFVI):** a set of physical hardware and software resources building the environment to host and connect the VNFs. NFVI creates a virtualized environment for all the VNFs allowing them to communicate with the MANO via the Nf-Vi connection point.

2.2. 5G Service Based Architecture

The 3GPP standardization body introduces the 5G core network architecture in the technical specification (TS 23.501) as service-based architecture (SBA), where the elementary module is the "Service" which defines a Network Function (NF). In SBA the architectural elements defined as NFs offer their services via interfaces of a common framework to any NFs that are allowed to make use of these provided services. The SBA realizes the flexible addition and extension of functions including proprietary ones leveraging to connect to other components without introducing specific new interfaces. The 5G SBA specified by the 3GPP is shown in Figure 1 where several NFs are defined, such as SMF, AMF, UPF, etc. The Authentication Server Function (AUSF) uses the Extensible Authentication Protocol (EAP) for authentication; the Access and Mobility Management Function (AMF) is responsible for hosting all mobility management-related functionality and terminates the non-access stratum (NAS) and N2 reference point messages. The NAS messages are sent between the UE and the AMF for mobility management and between the UE and Session Management Function (SMF) for session management. The Network Slice Selection Function (NSSF) selects the set of network slice instances to serve the UE; the Policy Control Function (PCF) provides policy rules to the control pane functions; and the SMF controls the User Plane Function (UPF) to manage the session (e.g., session establishment, modify, and release). The Unified Data Management (UDM) stores the subscription and authentication data, and the UPF is the user plane gateway serving the UE by connecting the RAN nodes to the data network (DN). The Application Function (AF) enables dynamic policy and charging control for applications. The UE is the mobile terminal and the DN may be operator services, internet access, or third-party services [15].

Under this architecture, real-time communication between each NF is greatly guaranteed through the central service bus connection and the communication path of each NF service can be optimized according to the demand. The communication between the NF service is realized through REST programming application interface (API) calls using a standard HTTP/2 routing mechanism. The HTTP/2 over Transmission Control Protocol (TCP) is used as the protocol to invoke the API(s). However, the access elements in the network such as Radio Access Network (RAN), UE, and NFs transferring the user plane communication still use the dedicated reference points.

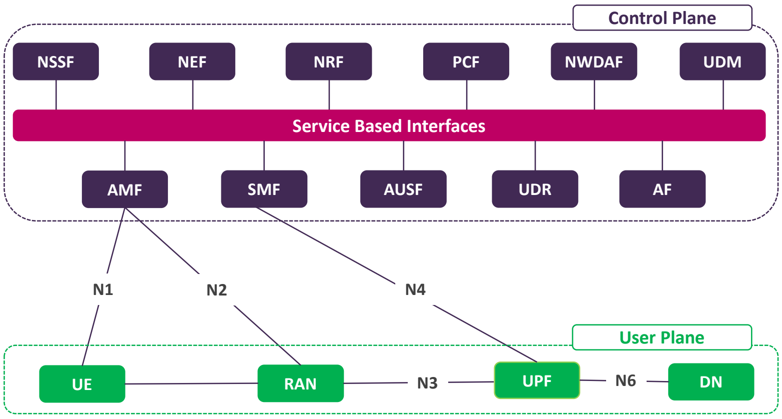

Figure 1. 5G Service Based Architecture.

The three main components of an SBA service framework are service registration, discovery, and authentication. To execute the service framework, the Network Repository Function (NRF) is introduced. A new service registers with NRF by providing the relevant information such as the type of service and network address characteristics etc. In this way, the service acts as a service provider, allowing it to be discovered anytime it is needed. Service consumers are those who use a particular service. Therefore, the service discovery method allows a service consumer to find the right service provider while the authentication system ensures that only those who have been granted access to the service can use it [16].

2.3. Network Data Analytics Function

The 3GPP (TS 29.520) introduces relatively a new Network Data Analytics Function (NWDAF) in the 5G SBA as a core network function that supports intelligent and autonomous network operations and service management. NWDAF is a general architecture through which various data-driven artificial intelligence (AI) / machine learning (ML) analytical technologies can be integrated with the 5G networks. Figure 2 shows the NWDAF architecture where it collects the data from different modules such as network functions e.g., (AMF, SMF, and PCF) application functions (AFs), the unified data repository (UDR), and operation, administration, and management (OAM) system to perform the network analytics.

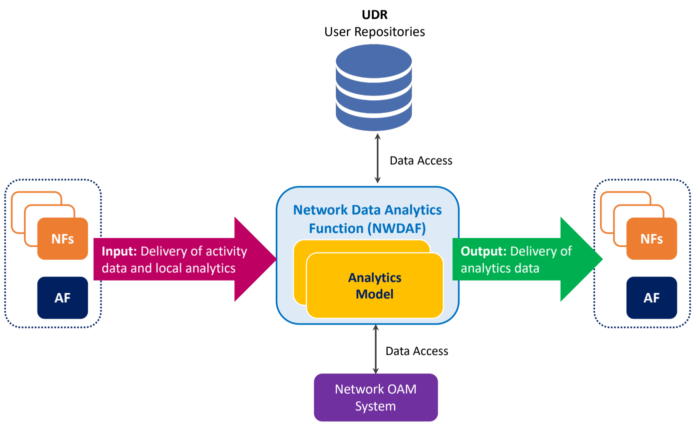

Figure 2. Architecture of Network Data Analytics Function.

Typically, the data are collected from the NFs when the data are related to the individual UE or network sessions. However, the data related to the global network states, such as performance measurements of network slices and services/applications, are often obtained from OAM. NWDAF can collect the data in two ways either reactively requests the particular customer (NF/AF/OAM) for a certain type of information or proactively retrieves the data for long-run analytics. The NWDAF realizes the data collection via a service-based interface (SBI) using either the request/response mode to collect the data for one time or subscription/notification mode to regularly retrieve the data for a certain time period. Once the data are collected, it performs the analytics and delivers this analytics information to the requested NFs, AF, and OAM to make optimal decisions about network operations and management actions. The NWDAF generates two main types of analytical information (i) statistical information and (ii) prediction information [17]. The statistical information includes traffic load and resource utilization of network slices, network/service performance measurements, and user mobility pattern analysis. However, the prediction information comprises temporal and spatial traffic distribution predicted for a future time period and UE location prediction for a future time instant. The 5G SBA allows the NWDAF to be deployed as single or multiple virtual network functions (VNF) on the same network domain. The multiple instances of NWDAF can be deployed as either a central NF, a collection of distributed NFs, or a combination of both. When multiple instances of NWDAF are deployed, each instance may provide a certain type of analytics. The interaction between the NWDAF and the NFs is a consumer-provider model. That is, 5G SBA allows the consumer of the NWDAF to discover and select the suitable NWDAF instance to obtain the specific analytics via the network NRF. After discovering an NWDAF instance, the consumer NFs can either retrieve the data via SBI request/response messaging protocol to obtain one-time analytical information per request or may subscribe to a certain type of analytic service offered by the NWDAF instance to regularly receive notification messages from the NWDAF instance that delivers analytical information generated from the service.

Concludingly, some of the main use cases of 5G NWDAF are: (i) load-level computation and prediction for a network slice instance; (ii) load analytics information and prediction for a specific NF; (iii) network load performance computation and future load prediction; (iv) UE mobility-related information and prediction; (v) QoS sustainability which involves the reporting and prediction of QoS change [18].

3. Integration of the Enabling Technologies

In this section, we highlight the need for integration of the various enabling technologies and concepts that are instrumental in realizing the concept of network slicing by exploiting the architectural components of 5G's SBA. We believe an analysis of the renowned 5G platforms would enable the readers to not only comprehend the said integration but also understand the differentiating features of each platform. Furthermore, we also believe that the contents of this section will provide a quick hands-on start for the researchers in this domain. In what follows next, we introduce and compare various 5G RAN, Core and Network Functions Virtualization Orchestrator platforms carefully. The comparison is provided in Tables 1–3 respectively. However, the readers are encouraged to look into the platforms' documentation for more details.

3.1. Radio Access Network Platforms

The comparison between the RAN solutions/platforms is based on eight factors such as the support for eNodeB (eNB), gNodeB (gNB), Multiple Input Multiple Output (MIMO), Software Defined Radio (SDR) user equipment, Commercial Off-the-Shelf (COTS) user equipment, type of license, main contributors, and type of community support. The comparison shown in Table 1 shows that the OAI RAN and srsLTE provide both the implementation of eNB and gNB whereas the Radisys RAN only provides the implementation of gNB. Radisys contributes to O-RAN providing an open-source implementation of the 3GPP new radio stack for the gNB distributed unit. O-RAN is open, intelligent, and deployed on a virtualized platform with high flexibility. The concept behind the O-RAN is all about disaggregating the hardware from software: open hardware and software using the standard processors with open interfaces. All RANs, listed in the table support the MIMO technology for the uplink and downlink. Furthermore, the OAI RAN and srsLTE can be used to serve both COTS and SDR UEs, however, this information is not available for the Radisys RAN. In terms of the license, the OAI RAN is distributed under the Apache License v1.1 which enables individuals and companies who have patents to contribute to the OAI source code while maintaining their patent rights. The Radisys is distributed under the Apache v2.0 license and the srsLTE is written in C++ and released under the GNU AGPLv3 license.

Table 1. Comparison of Radio Access Network Platforms.

Category	Platforms/ Framework	Ref.	Features							
			eNB	gNB	Support for MIMO	Support for SDR UE	Support for COTS UE	Licence	Main Contributor	Community Support
Radio Access Network	OAI RAN	[19]	Yes	Yes	Yes	Yes	Yes	OAI public license v1.1	OAI software alliance, EURECOM	Mailing List
	Radisys open source RAN	[20]	No	Yes, (O-RAN)	Yes	-	N/A	Apache v2.0, O-RAN Software License v1.0	Radisys	No
	srsLTE	[21]	Yes	Yes	Yes	Yes	Yes	GNU AGPLv3	Software radio systems	Mailing List

3.2. Core Network Platforms

In this sub-section, the main solutions to deploy the core network are discussed briefly. The comparison between the core solutions/platforms is based on seven factors such as the operational mode, support for core functions/components, supported interfaces, network slice, type of license, main contributors, and type of community support. The comparison shown in Table 2 shows that OAI-CN and free5GC projects focus only on the development of the 5G Standalone (SA) operational mode, however, the Open5GS, Open5GCore, SD-Core, Magma Core, and Nokia 5G core implement both 5G SA and 5G Non-Standalone (NSA) modes. In addition to this, some of these platforms also provide support for 4G, LTE, and NB-IoT LTE modes. In terms of network functions, most of the platforms can deploy AMF, SMF, AUSF, UDM, UDR, NRF, etc functions. However, OAI-CN, free5GS, Open5GCore, and Nokia 5G Core platforms can also deploy the newly introduced Non-3GPP Interworking Function (N3WIF) which is responsible to interwork between untrusted non-3GPP networks and the 5G Core. It acts as a gateway for the 5G core network providing the support for N2 and N3 interface towards the 5G core. Additionally, N3IWF also provides a secure connection for the user equipment accessing the 5G core over a non-3GPP access network with support for internet protocol security (IPsec) between the UE and the N3IWF. Magma Core can only deploy the AMF function since the remaining functions are still under development. Considering the importance of intelligence in 5G core networks, it is important to highlight that only the Nokia 5G Core platform can deploy the NWDAF function which supports intelligent and autonomous network operations and service management. On the other hand, most of the platforms listed in Table 2 bring great convenience to the implementation of network slicing such as OAI-CN, free5GC, Open5GCore, etc. Finally, in terms of the license, the OAI-CN, free5GC, and COMAC are distributed under the Apache V2.0 license, Open5GS is under the GNU AGPL3 license, Magma Core is under Berkeley Source Distribution (BSD) license, and the SD-Core is released with an ONF member-only license.

Table 2. Comparison of 5G Core Network Platforms and Framework.

Category	Platforms/ Framework	Ref.	Features							
			Operational Mode	Functions/ Components	Supported Interfaces	Network Slice Support	Supported Operating System	License	Main Contributors	Community Support
Core Network	OAI-CN	[22]	5G SA	AMF, SMF, NRF, AUSF, UDM, UDR, N3IWF, and UPF	-	Yes	Ubuntu 14.04 LTS and Ubuntu 16.04	Apache v2.0	OpenAirInterface software alliance, EURECOM	Mailing List
	free5GC	[23]	5G SA	AMF, SMF,UPF, OAM, N3IWF, UDM	N1/N2, N3/N4/N6 N8, N10/N11, N12/N13	Yes	Ubuntu 18.04	Apache v2.0	Free5GC	Forum
	Open5GS	[24]	5G NSA Core, 5G SA	AMF, SMF, AUSF, UDM, NRF, UPF	-	-	Ubuntu, openSUSE, CentOS, Fedora, and Mac OS	GNU AGPLv3	Open5GS	Mailing List/ Forum
	Open5GCore	[25]	5G SA, 5G NSA, LTE, NB-IoT LTE	AMF, SMF, AUSF, UDM, NRF, UPF, UDR, BSF, N3IWF	5G Interfaces (N1, N2, N3, N4)	Yes	-	Paid License	Fraunhofer FOKUS	Tutorials and training sessions
	SD-Core	[26]	LTE, 5G NSA and 5G SA	AMF, SMF, AUSF, PCF, NRF, UDM, UDR, NSSF	-	Yes	-	Member-Only Software License	ONF	Mailing List
	Magma Core	[27]	4G, NB-IoT Support, 5G NSA, 5G SA	AMF * Remaining functions are under development	N1,N2,N3	Yes	Mac OS, Ubuntu 20.04	BSD	Facebook	Mailing List/ Forum
	COMAC	[28]	-	-	-	-	-	Apache v2.0	ONF	Mailing List
	Nokia 5G Core	[29]	5G NSA, 5G SA	UPF, N3IWF, NEF, NRF, NWDAF, PCF, SMF, AUSF, UDM, UDR	-	Yes	-	-	Nokia	Training

3.3. Network Functions Virtualization Orchestrator Platforms

In this sub-section, the well-known platforms to deploy NFV Orchestration are discussed briefly. The comparison between the platforms is based on nine factors, such as Compliance with ETSI MANO architecture, support for external APIs, type of network services, components/frameworks, infrastructure, supported VIM types, community, and installation prerequisites. The comparison shown in Table 3 shows that the MANO OSM, Cloudify, and Open Baton are fully-compliant with the ETSI NFV MANO architecture, however, the ONAP provides partial compliance with this architecture. The OSM community is compliant with the MANO reference architecture, following the IFA working group specifications. Key reference points, such as Or-Vnfm and Or-Vi might be identified within OSM components. Furthermore, the VNF and network service catalogue are explicitly available in an OSM service orchestrator component. ONAP, OSM, and Cloudify frameworks are equipped with APIs offered to OSS/BSS and other relevant services, however, the Open Baton provides access to Java SDK external API. In terms of network service and support for network slicing; all platforms listed in the Table provide the VNFs service and the ONAP, OSM, and Open Baton also support network slicing at the infrastructure level. All these frameworks leverage the virtual machine, and containers with Kubernetes, and Docker to deploy the platforms. Some of the other comparison factors along with platform installation prerequisites are discussed in Table 3.

Table 3. Comparison of NFV Orchestration Platforms.

Category	Platforms	Ref.	Features						Minimum Installation Prerequisites					
			Compliance w/ETSI MANO	External APIs	Network Service	Support for Network Slicing	Components/ Frameworks	Infrastructure	Supported VIM Types	Community	CPUs/ vCPUs	RAM	Storage	Supported Operating System
Virtual Network Function Orchestrator	ONAP	[30]	Partially	REST APIs (for external controllers, OSS/BSS, etc.)	VNFs, PNFs	Yes	(i) Management Framework (ii) Design Time Framework (iii) Run Time Framework	Virtual machine, Containers w/Kubernetes & Docker	OpenStack Kubernetes Vmware Azure	Linux foundation w/telecom operators	112 vCPUs	224 GB	160 GB	Ubuntu 14.04 Ubuntu 16.04
	MANO OSM	[31]	Yes	REST APIs (for external controllers, OSS/BSS, etc.)	VNFs, PNFs	Yes	(i) Information Model (ii) OSM Automation Framework	Virtual machine, Containers w/Kubernetes	OpenStack Vmware OpenVIM Kubernetes	ETSI w/telecom operators	2 CPUs	6 GB	40 GB	Ubuntu 20.04
	Cloudify	[32]	Yes	Asynchronous RESTful API/Users & Services authorization	VNF	N/A	-	VMs/Containers	OpenStack	Cloudify Team	2 vCPUs	4 GB	5 GB	Ubuntu 14.04 CentOS 7.1 OS X 10.11
	Open Baton	[33]	Yes	Java SDK	VNFs	Yes	-	Virtual machines, Containers w/Docker	OpenStack	Fraunhofer FOKUS & TU Berlin	2 CPUs	2 GB	10 GB	Ubuntu 14.04 Debian Jessy

4. Potential Way Forward to 6G

The stakeholders are figuring out 6G design objectives as 5G is nearly at the point of technological realization in all of its features. To successfully operate internet of everything

(IoE) services such as extended reality (XR) and networked autonomous systems, a wireless system must concurrently deliver high reliability, low latency, and high data rates, for heterogeneous devices, across uplink and downlink. Additionally, end-to-end co-design of computation, control, and communication functionalities will be necessary for emerging IoE services, which have been mostly neglected up until now. To support this new class of services, specific challenges must be overcome, such as characterizing the fundamental rate-reliability-latency tradeoffs that govern their performance, exploiting frequencies above sub-6 GHz, and converting wireless systems into a self-sustaining, intelligent network fabric that flexibly provisions and orchestrates communication-computing control-localization-sensing resources tailored to the required IoE scenario [34].

The various new use cases that we anticipate by 2030 and beyond will determine the new standards that 6G must meet. The data rate/throughput/capacity, latency, reliability, scale, and flexibility key performance indicators (KPIs) from 5G will continue to be crucial benchmarks for 6G performance. For 6G's anticipated use cases, several additional features will also be crucial, including (i) dynamic digital twins and virtual worlds; (ii) zero-energy devices; and (iii) wireless in data centres, etc [10]. In Figure 3, we categorise the requirements for 6G into six categories three of which have KPIs identical to those of 5G and three of which are the new ones.

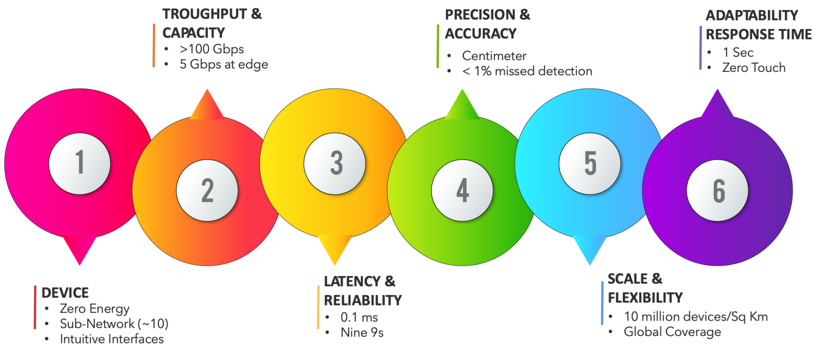

Figure 3. Key requirements and characteristics of 6G.

6G is also anticipated to address the difficulties unique to a space-air-ground integrated network (SAGIN). An integrated mobile communication network called SAGIN combines satellites, airborne platforms and terrestrial networks. The terrestrial networks are capable to provide standardized coverage of urban hotspots. In remote locations, on the sea, and in the air, satellite and airborne technology provide ubiquitous coverage. SAGIN is a deep integration of a system, technology, and application rather than merely connecting various communication networks. By integrating the RAN and CN, SAGIN creates a unified control plane and data plane, as well as a uniform terminal and air interface protocol. A robust control plane is developed by leveraging the deployment of the control network element. Moreover, to achieve seamless handover, intelligent mobility management technology is also deployed. Differentiated micro-service network elements with space-air characteristics are used to accomplish the space-air QoS provisioning, policy routing, resource allocation, and other network operation management activities.

Despite the conceptual infancy of SAGIN, relevant stakeholders have provided several important insights into the improvement of terrestrial, aerial, and satellite systems. SAGINs are anticipated to cover a wider range of smart city scenarios, which present enormous and constantly changing diversified service requirements. This encourages the development of the IMT2020 usage case scenarios into further-eMBB (feMBB), ultra-mMTC (umMTC), extremely reliable and low latency communications (ERLLC), mobile broadband reliable low latency communications (MBRLLC), and massive-uRLLC (muRLLC). This is to say that factors such as network resource scheduling, round trip duration, throughput, handover management, etc. are essential for a higher level of autonomous driving. Even though centralized broadcast scheduling guarantees excellent broadcast reliability, it is burdened by signalling overheads caused by dynamically changing vehicle positioning and network resource demands. Therefore, the distribution version of resource scheduling implanted with autonomous behaviour provides improved scalability, which is also reflected in enhanced sidelink transmission, carrier aggregation, 5GNR, support for unicast and multicast, etc. of 3GPP Release 15 and 16 [35]. The improvement in URLLC, positioning precision, and reliability over 6G is further aided by 3GPP Release 17. The pervasive network intelligence made possible by 6G is anticipated to offer greater service diversity and flexible service configuration. AI-enabled solutions for service composition can address issues with service-oriented networks, such as huge data exploitation through the use of big data approaches and service identification through intent-driven networking.

Zhang et al. [36] investigate the support for vehicle networks in the envisioned integrated space-air-ground and emphasize the need for integrated communication to support the requirements of various vehicular services, particularly in cases of frequent handovers. The authors highlighted a variety of challenges preventing the implementation of such integrated communication systems, including those of heterogeneous management, dynamic networking, service quality, etc. In another, research the research literature survey of 6G SAGINs from a service-oriented network perspective was carried out by Cheng et al. [37]. The authors provide a brief overview of the various SAGIN types, which include Cooperative satellite-terrestrial networks, Cognitive satellite-terrestrial networks, Hybrid satellite-terrestrial relay networks, and Satellite-terrestrial Backhaul Networks. The advantages of SAGINs for autonomous driving were also underlined by the authors, including LEO's low cost and ubiquitous connection, the wide carrier coverage that reduces handovers, caching of the pertinent contents in UAVs, etc.

5. Dynamic Network Slicing—A Dilemma for Highly Changing Environment

Now that the reader is equipped with the necessary background and platforms, in this section, we highlight a few crucial challenges of network slicing that this study addresses:

- *Challenge 1—Dynamically adapting application service quality requirements:* As the world transits to adopting new solutions in different verticals e.g., remote surgery in eHealth, autonomous driving in transportation, smart homes, smart farms, etc, the service quality for network relevant service are subject to frequently changing owing to highly dynamic environments. Hence, a network slice for a longer period may not fall under optimal solutions.
- *Challenge 2—Unpredictable Resource and Service Quality Demands:* The classical method of resource allocation, bearer formation (e.g., in LTE), and network slicing are driven by demand estimation, which has been relatively deterministic owing to the type of services being used so far. However, services of the future are expected to ask for dynamically changing network slices on shorter time quanta, which is usually the consequence of a fast-changing environment. The obvious challenge here is the accurate demand estimation.

To address the above challenges, this study proposes the concept of a data-driven proactive network slicing mechanism detailed below.

This research aims to build a predictive model to predict the optimal network slice for a requested service resultantly assisting in allocating enough resources to that slice based on the traffic prediction. The big picture of data-driven network slice prediction is shown in Figure 4. The proposed methodology is comprised of the following steps; (i) the network slice data are downloaded, provided by Anurag et al. [38]; (ii) the data are pre-processed and cleaned to build models; (iii) the cleaned dataset is then, divided into training and testing sets; (iv) the predictive models are developed for three machine learning algorithms to predict the network slice type; (v) testing data are then given to the model to test the model performance; (vi) finally, the results of all the algorithms are evaluated and compared.

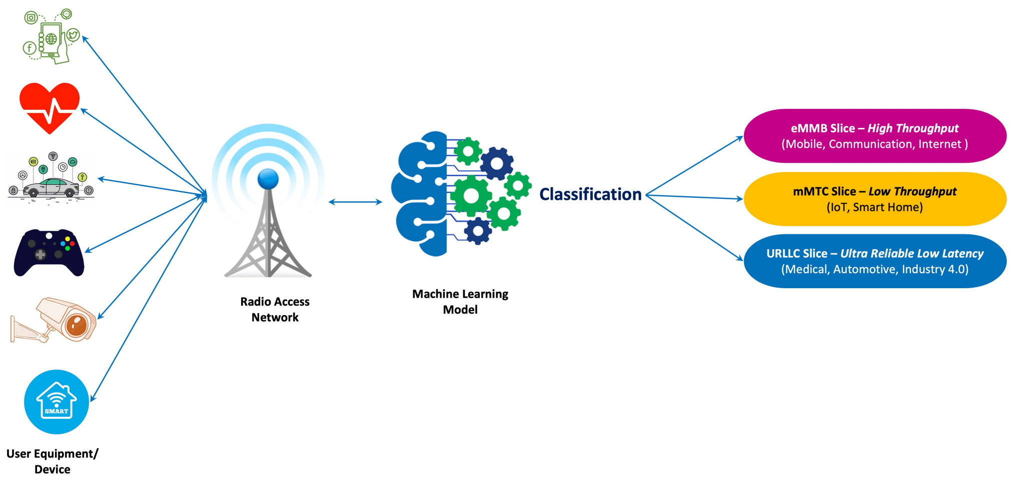

Figure 4. Big picture of Data-Driven Network Slice Prediction.

5.1. Dataset

In this research, the DeepSlice dataset published by Anurag et al. [38] is used to predict the optimal network slice for all incoming requests. The dataset is comprised of 63,168 instances and 8 features. Some of the main features and their possible values are discussed in Table 4.

Table 4. An Overview of DeepSlice Dataset.

Feature	**Values**
Use case Type	Smartphone, IoT Devices, Smart Transportation, Industry 4.0, AR/VR/Gaming, Healthcare, Smart City/Home
User Equipement Category	LTE, 5G
Technology Supported	LTE/5G, IoT (LTE-M, NB-IoT)
Guaranteed Bit Rate (GBR)	GBR, Non-GBR
Reliability	0.01, 0.001, 0.000001
Latency (Ms)	10, 50, 60, 75, 100, 150, 300

5.2. Preprocessing

Data preprocessing and cleaning are the most important stages to handle the data before using them in machine learning algorithms. Without preprocessing, the raw data cannot be transformed into a useful and efficient format to create a machine-learning model. The DeepSlice dataset is available in CSV format. After downloading the dataset, we remove the duplicate instances. After then, we perform the attribute value transformation from string to numeric values.

5.3. Data Splitting

The cleaned dataset is then split into 80% training and 20% testing sets. To keep the target class "slice type" balanced, its distribution should be the same in both training and testing datasets ensuring a more balanced dataset instead of a skewed one. Therefore, we equally distributed the data.

5.4. Machine Learning Algorithms for Slice Prediction

The training data (80%) is given to machine learning algorithms to train the model. We build the classification models, and the slice type attribute is used as the target variable. Three machine learning algorithms: Random Forest (RF), eXtreme Gradient Boosting (XGBoost) and Neural Network (NN) are trained, and compared to classify the slice type of a request into eMBB (enhanced Mobile Broadband), URLLC (Ultra Reliable Low Latency Communications), mMTC (massive Machine Type Communications), and master slice based on the following parameters such as throughput, latency, equipment type, reliability, mobility, etc. RF is a supervised learning model mostly used to construct predictive models for both regression and classification problems. The reason behind choosing RF over the other algorithms such as Decision Tree or Naïve Bayes is that we leveraged the well-structured data, therefore, the RF uses several sub-trees to lessen the chance of overfitting. The RF works well with large datasets and outputs accurate predictions. Moreover, it is capable of determining the missing values and maintaining the accuracy of the model even if some input data are missing. XGBoost is a branch of Gradient Boosting Machine (GBM) techniques used to build both classification and regression predictive models. XGBoost is an ensemble approach, where new models are built to eliminate residuals or other errors from earlier models before being integrated to provide the final prediction. Experimentally, XGBoost outperforms many ensemble classifiers in terms of speed and performance. Lastly, a three-layer feed-forward neural network (FNN) with a ReLu activation function is implemented to predict the slice. The selected parameters used to train the classifiers are shown in Table 5. After training the classifiers the testing data (30%) are given to the model to predict the slice type.

Table 5. Selected Parameters for Algorithms.

Algorithm	Parameters
Random Forest	Batch size = 100; Iterations = 100; Seed = 1
XGBoost	Default Parameters
Neural Network	Epochs = 5; Batch size= 16; Learning Rate = 0.003 Dropout = (p = 0.2); Optimizer = Adam; Loss = CrossEntropyLoss

5.5. Experimental Results and Discussion

The experiments are conducted on Windows 10 operating system and Python version 3.7. The multi-class confusion metric [39] is used to evaluate the performance of the algorithms where the experimental results show that both RF and XGBoost give significantly good accuracy to predict the slice for all incoming requests to the network. The logic behind the random forest is the bagging principle, which improves the overall performance by adding more randomness to the model as the trees grow, which; resultantly contributes to achieving high accuracy. Besides the good performance of random forest, it also takes less time to build and run the model.

Interpretability refers to the ability to explain what the model is doing; it is a crucial step in the ML pipeline. Therefore, to increase the understanding of our model we leveraged the feature importance concept. Another reason to choose the feature importance is to help us to determine whether the predictions made by the model are sensible. Figure 5 and 6 show the feature importance of both algorithms. The feature importance is referred to as an assessment of the specific contribution made by the relevant feature for a given classifier regardless of the shape or orientation of the feature effect. This is to say that the features of the input data have varying degrees of importance depending on the classification model used and that a feature that is significant for one model may not be significant for another [40]. Therefore, it is imperative to determine the degree of usefulness of a specific variable for a current model and prediction. Figure 5 clearly shows that, with RF, the technology supported, latency, use case type, and reliability features are more significant.

On the other hand, in Figure 6, XGBoost gives more significance to the use case type, GBR, latency and technology-supported features respectively to train and test the model. These results indicate that GBR, latency and reliability requirements play an important role in accurately inferring the slice types. Furthermore, the Figures highlight interesting insight that both algorithms assign zero importance to day and time features showing no importance of these features in training the models.

For ML classification problems a state-of-the-art way to evaluate the model's performance is to compute its accuracy. Therefore, to further evaluate the effectiveness of the proposed model we leveraged the accuracy metric. In the proposed research the accuracy aims to determine the capability of the model to accurately predict the slice type. Figure 7 shows the training and testing accuracy of the neural network and it is evident from the result that the model achieved significant accuracy.

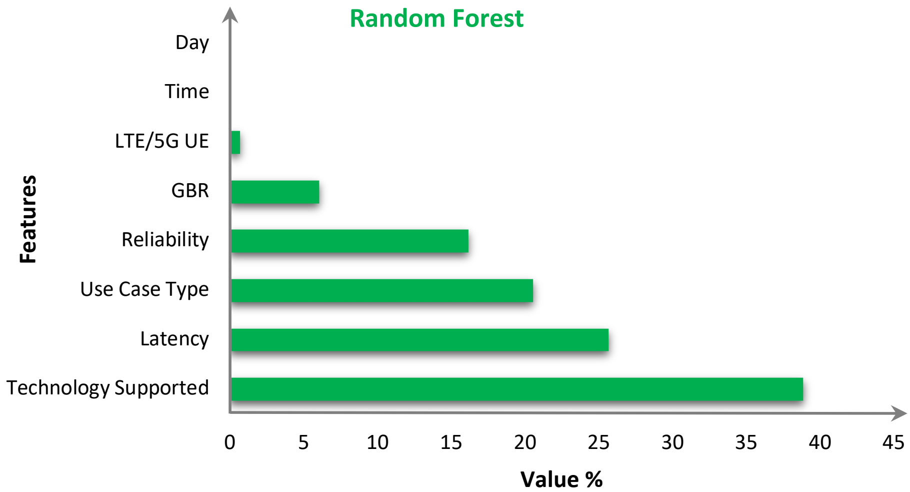

Figure 5. Feature Importance of Random Forest Algorithm.

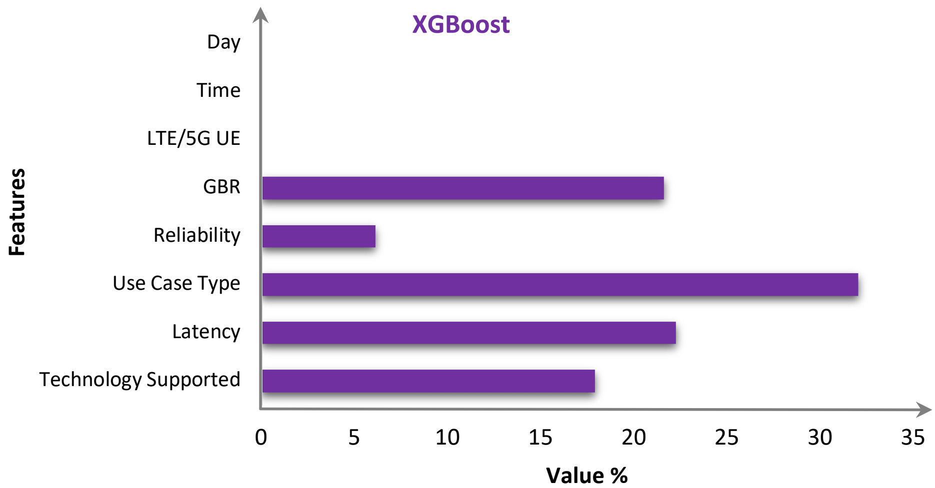

Figure 6. Feature Importance of XGBoost Algorithm.

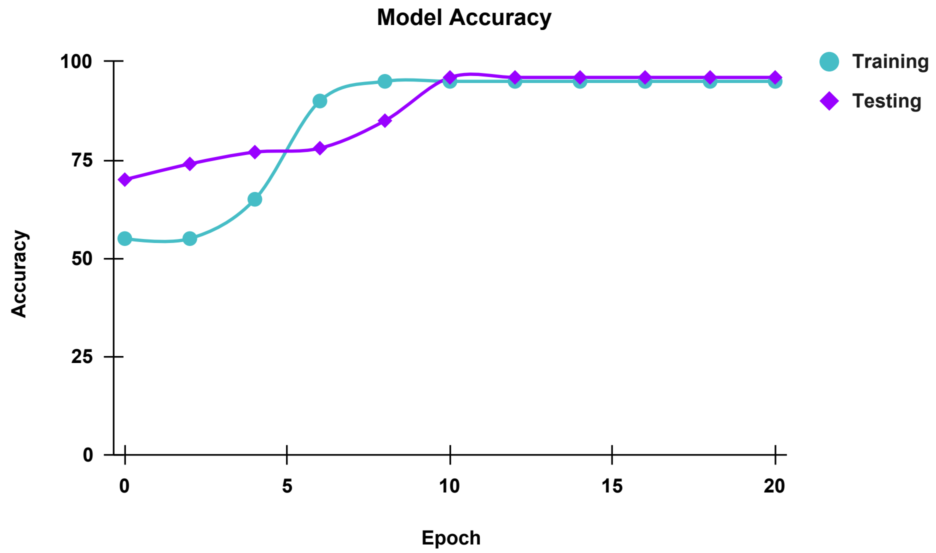

Figure 7. Training and validation accuracy of Neural Network.

6. Conclusions and Future Work

This paper is mainly divided into four parts. The first part provides an overview of 5G enabling technologies and network slicing to comprehend the content of this paper. The second part highlights the need to integrate the enabling technologies that are important in realizing the concept of network slicing by exploiting the architectural components of 5G's SBA. Furthermore, it equips the reader with an overview of 5G platforms. The fourth section discusses the potential ways towards 6G. The fifth section provides the methodology and the results of the initial deployment of 5G network slicing leveraging the machine learning algorithms. A predictive classifier is developed to predict the slice for each incoming request to the network. Three machine algorithms (random forest, XGBoost and neural network) are used to deploy the model and the experiments show significantly promising results. The limitation of the proposed work is that the model is trained on a set of data which may be changed in real-time scenarios, hence the accuracy may fail. We are digitalizing the roads of the United Arab Emirates University making them smarter by deploying smart towers equipped with heterogeneous sensors, communication infrastructure, and AI toolbox. In this connection, the proposed framework will be used to deploy the 5G core to implement various use cases such as autonomous driving and ehealth. Moreover, we also plan to completely automate the proposed network slicing framework by leveraging the NWDAF function and the 5G platforms discussed in the paper.

Author Contributions: Conceptualization, S.M., M.A.K., A., Investigation, S.M., M.A.K., A. and H.E.-S.; Methodology, S.M., M.A.K. and A., H.E.-S.; Project administration, S.M., J.K. and O.U.; Supervision, M.A.K. and H.E.-S.; Writing—original draft, S.M.; M.A.K.; Writing—review and editing, S.M., M.A.K., A., H.E.-S., J.K. and O.U. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Emirates Center for Mobility Research (ECMR) of the United Arab Emirates University (grant number 31R151).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

3GPP	3rd Generation Partnership Project
5G	Fifth Generation
6G	Sixth Generation
AF	Application Function
AI	Artificial Intelligence
AMF	Access and Mobility Function
AUSF	Authentication Server Function
BSS	Business Support Systems
COTS	Commercial-Off-The-Shelf
DN	Data Network
eMBB	enhanced Mobile Broadband
eNB	eNodeB
ETSI	European Telecommunications Standards Institute
FNN	Feedforward Neural Network
gNB	gNodeB
HTTP/2	Hypertext Transfer Protocol (Second Version)
LTE	Long Term Evolution
MANO	Management and Orchestrator
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
mMTC	massive Machine Type Communications
N3WIF	Non-3GPP Interworking Function
NB-IoT	NarrowBand-Internet of Things
NFVI	Network Function Virtualization Infrastructure
NF	Network Function
NFV	Network Function Virtualization
NRF	Network Repository Function
NSA	Non-Standalone
NSSF	Network Slice Selection Function
NWDAF	Network Data Analytics Function
O-RAN	Open Radio Access Network
OAM	Operation, Administration, and Management
ONF	Open Networking Foundation
OSS	Operations Support Systems
PCF	Policy Control Function
QoS	Quality of Service
RAN	Radio Access Network
RF	Random Forest
SDR	Software-Defined Radio
SBA	Service Based Architecture
SBI	Service Based Interface
SMF	Session Management Function
SA	Standalone
UDR	Unified Data Repository
UE	User Equipment
URLLC	Ultra Reliable Low Latency Communications
VNF	Virtual Network Function

References

1. Alliance, N. NGMN 5G White Paper v1. 0. Approved and Delivered by the NGMN Board, 17 Feb 2015. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4003830 (accessed on 25 July 2022).
2. Shen, X.; Gao, J.; Wu, W.; Lyu, K.; Li, M.; Zhuang, W.; Li, X.; Rao, J. AI-assisted network-slicing based next-generation wireless networks. *IEEE Open J. Veh. Technol.* **2020**, *1*, 45–66.
3. Chen, W.K.; Liu, Y.F.; De Domenico, A.; Luo, Z.Q.; Dai, Y.H. Optimal network slicing for service-oriented networks with flexible routing and guaranteed E2E latency. *IEEE Trans. Netw. Serv. Manag.* **2021**, *18*, 4337–4352.

4. An, N.; Kim, Y.; Park, J.; Kwon, D.H.; Lim, H. Slice management for quality of service differentiation in wireless network slicing. *Sensors* **2019**, *19*, 2745.
5. Dong, T.; Zhuang, Z.; Qi, Q.; Wang, J.; Sun, H.; Yu, F.R.; Sun, T.; Zhou, C.; Liao, J. Intelligent joint network slicing and routing via GCN-powered multi-task deep reinforcement learning. *IEEE Trans. Cogn. Commun. Netw.* **2021**, 1269–1286
6. Ebrahimnejad, A.; Karimnejad, Z.; Alrezaamiri, H. Particle swarm optimisation algorithm for solving shortest path problems with mixed fuzzy arc weights. *Int. J. Appl. Decis. Sci.* **2015**, *8*, 203–222.
7. Di Caprio, D.; Ebrahimnejad, A.; Alrezaamiri, H.; Santos-Arteaga, F.J. A novel ant colony algorithm for solving shortest path problems with fuzzy arc weights. *Alex. Eng. J.* **2022**, *61*, 3403–3415.
8. Abbaszadeh Sori, A.; Ebrahimnejad, A.; Motameni, H. Elite artificial bees' colony algorithm to solve robot's fuzzy constrained routing problem. *Comput. Intell.* **2020**, *36*, 659–681.
9. Sori, A.A.; Ebrahimnejad, A.; Motameni, H.; Verdegay, J.L. Fuzzy constrained shortest path problem for location-based online services. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **2021**, *29*, 231–248.
10. Saad, W.; Bennis, M.; Chen, M. A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. *IEEE Netw.* **2019**, *34*, 134–142.
11. Rani, S.J.; Ioannou, I.; Nagaradjane, P.; Christophorou, C.; Vassiliou, V. A Two-Stage Machine Learning Approach for 5G Mobile Network Augmentation through Dynamic Selection and Activation of UE-VBSs. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4003830 (accessed on 26 July 2022).
12. Chen, W.K.; Liu, Y.F.; Dai, Y.H.; Luo, Z.Q. Optimal Qos-Aware Network Slicing for Service-Oriented Networks with Flexible Routing. In Proceedings of the ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 22–27 May 2022; pp. 5288–5292.
13. Elsherif, F.; Chong, E.K.; Kim, J.H. Energy-efficient base station control framework for 5G cellular networks based on Markov decision process. *IEEE Trans. Veh. Technol.* **2019**, *68*, 9267–9279.
14. Virtualisation, N.F. An introduction, benefits, enablers, challenges & call for action. In Proceedings of the White Paper, SDN and OpenFlow World Congress, Darmstadt, Germany, 22–24 October 2012; p. 73.
15. Rudolph, H.C.; Kunz, A.; Iacono, L.L.; Nguyen, H.V. Security challenges of the 3GPP 5G service-based architecture. *IEEE Commun. Stand. Mag.* **2019**, *3*, 60–65.
16. Wang, X.; Sun, T.; Duan, X.; Wang, D.; Li, Y.; Zhao, M.; Tian, Z. Holistic service-based architecture for space-air-ground integrated network for 5G-advanced and beyond. *China Commun.* **2022**, *19*, 14–28.
17. Duan, Q. Intelligent and Autonomous Management in Cloud-Native Future Networks—A Survey on Related Standards from an Architectural Perspective. *Future Internet* **2021**, *13*, 42.
18. Sevgican, S.; Turan, M.; Gökarslan, K.; Yilmaz, H.B.; Tugcu, T. Intelligent network data analytics function in 5G cellular networks using machine learning. *J. Commun. Netw.* **2020**, *22*, 269–280.
19. 5G RAN–OpenAirInterface. Available online: https://openairinterface.org/oai-5g-ran-project/ (accessed on 22 June 2022).
20. Open RAN Radisys. Available online: https://www.radisys.com/solutions/openran (accessed on 22 June 2022).
21. The srsLTE Project is Evolving. https://www.srslte.com/srslte-srsran (accessed on 22 June 2022).
22. 5G CORE NETWORK–OpenAirInterface. Available online: https://openairinterface.org/oai-5g-core-network-project/ (accessed on 22 June 2022).
23. free5GC. Available online: https://www.free5gc.org/ (accessed on 22 June 2022).
24. GitHub–open5gs/open5gs: Open5GS is a C-language Open Source implementation for 5G Core and EPC, i.e., the core network of LTE/NR network (Release-16). Available online: https://github.com/open5gs/open5gs (accessed on 22 June 2022).
25. Open5GCore Open5GCore. Available online: https://www.open5gcore.org/ (accessed on 22 June 2022).
26. SD-Core–Open Networking Foundation. Available online: https://opennetworking.org/sd-core/ (accessed on 22 June 2022).
27. Magma–Linux Foundation Project. Available online: https://magmacore.org/ (accessed on 22 June 2022).
28. COMAC–Open Networking Foundation. Available online: https://opennetworking.org/comac/ (accessed on 22 June 2022).
29. 5G Core (5GC): Creating What's Next Nokia. Available online: https://www.nokia.com/networks/5g-core/ (accessed on 22 June 2022).
30. Architecture–ONAP. Available online: https://www.onap.org/architecture (accessed on 22 June 2022).
31. OSM. Available online: https://osm.etsi.org/ (accessed on 22 June 2022).
32. Cloudify DevOps Automation & Orchestration Platform, Multi Cloud. Available online: https://cloudify.co/ (accessed on 22 June 2022).
33. Open Baton: An open source reference implementation of the ETSI Network Function Virtualization MANO specification. Available online: https://openbaton.github.io/ (accessed on 22 June 2022).
34. Viswanathan, H.; Mogensen, P.E. Communications in the 6G era. *IEEE Access* **2020**, *8*, 57063–57074.
35. Khan, M.A.; El Sayed, H.; Malik, S.; Zia, M.T.; Alkaabi, N.; Khan, J. A Journey towards Fully Autonomous Driving-Fueled by a Smart Communication System. *Veh. Commun.* **2022**, *36*, 100476.
36. Zhang, N.; Zhang, S.; Yang, P.; Alhussein, O.; Zhuang, W.; Shen, X.S. Software defined space-air-ground integrated vehicular networks: Challenges and solutions. *IEEE Commun. Mag.* **2017**, *55*, 101–109.
37. Cheng, N.; He, J.; Yin, Z.; Zhou, C.; Wu, H.; Lyu, F.; Zhou, H.; Shen, X. 6G service-oriented space-air-ground integrated network: A survey. *Chin. J. Aeronaut.* **2021 35, 9**.

38. Thantharate, A.; Paropkari, R.; Walunj, V.; Beard, C. DeepSlice: A deep learning approach towards an efficient and reliable network slicing in 5G networks. In Proceedings of the 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 10–12 October 2019; pp. 762–767.
39. Heydarian, M.; Doyle, T.E.; Samavi, R. MLCM: Multi-label confusion matrix. *IEEE Access* **2022**, *10*, 19083–19095.
40. Saarela, M.; Jauhiainen, S. Comparison of feature importance measures as explanations for classification models. *SN Appl. Sci.* **2021**, *3*, 1–12.