Next-Generation Telecom Infrastructure Automation: Leveraging AI And ML For Resilient, Self-Healing Networks

Shabrinath Motamary

Site Reliability Engineer 2, motamaryshabrinath@gmail.com,
ORCID ID: 0009-0009-6540-7585

With the increasing telecommunications demands and limited human and operational resources, telecom operators are turning toward AI and ML technologies to modernize their networks. Using these algorithms, operators can bring intelligence into the entire network, and thus acquire all the capabilities necessary there to allow self-healing and fault management automation at any level. As a consequence of this, key performance indicators (KPIs) and other vital scene information can be extracted from numerous operational and other source points in the network, gravitating to a cloud-based backbone. The coherent labelling and merging of these data can allow basic and advanced use cases based on knowledge extraction and pattern recognition algorithms. For either the limited classical network technology or the more complex modern optical ones, it is now demonstrated that modern AI algorithms can implement all important autonomic fault management functions: observability, controlled actuation, diagnosis, and actions.

The new automation platform is easily extensible to both new providers or operational technologies and application scenarios beyond the telecommunications scope. Moreover, this technology is poised to greatly simplify the increasingly complex operation of communication networks. The major drive behind the ongoing network automation effort is the knowledge extraction from and observability of the silent components and asset usage within the infrastructure. AI algorithms can solve the initial task of KPI extraction, pricing, and pooling of scene information within a cloud-based backbone solution platform. Classical techniques like rules-based analytical algorithms or renegade statistic approximators can extend the understanding of the observed network components. There are some first application scenarios of this kind emerging focusing on classic copper networks and cloud infrastructure monitoring or KPIs extraction using primary in-car equipment usage telemetry data or spectrum occupancy monitoring systems. More granular explanations of general network KPIs across the mean-line IT accountancy boundaries are expected there.

Keywords: Artificial Intelligence, automation, Fault Management, Machine learning, Network event classification, Network status forecasting, Resilience, Self-healing networks,Telecom Infrastructure Automation,AI in Telecom Networks,Machine Learning for Network Resilience,Self-Healing Networks,Autonomous Network Management,Intelligent Network Automation,Network Operations AI,Predictive Network Maintenance,Cognitive Network Management,AI-Driven Fault Detection,ML-Based Network Optimization,Zero-Touch Network Operations,Next-Gen Network Architecture,Resilient Telecom Systems,AI-Powered Network Recovery.

1. Introduction

Telecommunication plays a significant role to keep you connected with the world. Nowadays, OT environments and processes in the telecom industry are becoming modernized, intelligent and automated with excessive use of hardware and software. Telecom infrastructure market is anticipated to reach USD124.5 Billion by 2028, growing at a CAGR of 23.8 %. Telecom infrastructure is an always-on system which provides high availability and low-latency service. Fault can be created during upgradation, maintenance, component failure and even during normal operation. A business impact, due to faulty components, translates to signalling and voice unavailability leading to customer dissatisfaction. Regardless of the consequence, the automated fault resolution cycle in the viewpoint of ITIL4 takes a long time in comparison to the resolution time of the fault.

Traditional telecom infrastructure consists mostly of proprietary hardware like signalling gateway (SG), media gateway (MG), load balancers, fraud detection, firewalls, logging systems etc. with simple custom software. The monitoring and management system is built as a monolithic console (MCle) to give a view of the entire telecom network slice. MCle checks the health of each component/server, its resource consumption and operation with a periodic interval and checks per component, the alarms triggered by component os, application, network interfaces and logs etc. Since the hardware consists of proprietary and appliances designed using network processors and field programmable gate arrays (FPGAs), telemetry information collected from each hardware is not structured, it varies from component to component vendor and it is difficult to compose a single view of the entire setup due to diversity heterogeneity between components. Telecom industry is using ML to identify patterns for predictive risk management such as unavailability failure, CPU, memory clogging etc. in pre-deployment or off-line mode. ML also assists in churn detection and clustering to optimize resource allocation and capacity planning. Also, machine learning (ML) provides techniques to adapt to dynamic behavior of networks.

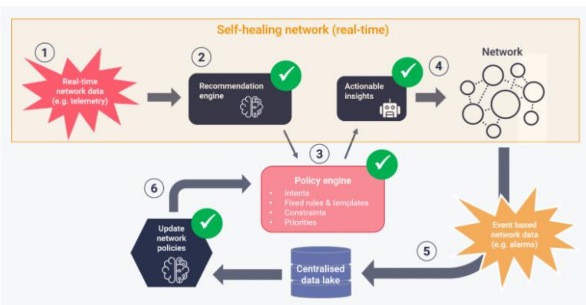

Fig 1: AI in network operations trust in a self-healing network

1.1. Background and significance

Flexible and less complex Telecom infrastructure, allowing for an easy roll out of services and products is the vision for telecom next generation infrastructure. Next generation telecom infrastructure will allow for Artificial Intelligence driven cognitive insights and decisions to allow for telecommunications network automation. This vision also includes new generation, more efficient products; flexible architecture to allow easy product roll out; and orchestration of 3rd party service products that will allow mutual service level agreements.

Telecom AI and ML provide a group of techniques to fundamentally adapt to the dynamic network behavior, predicting the future behavior, thus proactively deciding on mitigation, recovery, or adaptation actions. These AI and ML techniques and applications can help with network awareness, health, and capacity estimation. These learning-based techniques provide a promising platform for end-to-end network automation, including fault management.

Communication networks and services have become an indispensable part of modern society. However, their unprecedented complexity and rapid evolution make their correct operation a highly challenging task. Thus, the demand for advanced automation solutions is constantly increasing. Due to the enormous volumes of data and events of the order of tera-bytes generated in real-time and their complex interconnectivity, a set of techniques from the machine learning family have been gaining traction, namely: unsupervised learning, deep learning, reinforcement learning.

Equ 1: OSS/BSS System Autonomy Score (SAS)

$$SAS = \beta \cdot (1 - \frac{M_T}{T_O}) + \gamma \cdot L_{AI} + \delta \cdot R_{DevOps}$$

- M_T: Manual touchpoints
- T_O: Total operational tasks
- L_{AI}: Learning rate of AI models in real-time
- R_{DevOps}: Recovery rate from deployment failures
- β, γ, δ: Tunable weights based on domain priorities
- **Goal**: SAS approaching 1 indicates self-sufficient OSS/BSS

2. Background and Motivation

Telecommunications service providers (TSPs) worldwide strive to build networks able to handle unprecedented demand growth and complexity, while minimizing capital and operational expenditure (CAPEX/OPEX). Complexity originates not only from rapid growth in data traffic but also from more diverse services delivered across more heterogeneous networks. Recent surveys indicate networks already face a number of troubles: service and device failures, unscheduled maintenance, configuration errors, and performance issues.

A common characteristic found in all these failure types is triggered path errors. Devices fail mainly due to hardware malfunction or software failure. Scheduled maintenance results in devices being powered off or undergoing firmware updates, which alters their configuration. Configuration errors are usually introduced as part of deployment, and performance errors typically originate from automated probing test results or policies not taking in consideration all devices and services in a chain. Errors propagate along the network, producing a wide variety of negative effects such as degraded services, customer complaints, net loss of revenue.

Error management traditionally consists of devices logging alarms and a second operator looking into these log files, analyzing them manually, and taking appropriate action. This approach yields good results in a number of cases; however, it has a few limitations: it is resource hungry, incomplete, slow, and alert intuition-centric. Labor scarcity and increasing complexity drive the invention of automated solutions, able to speed up error management while reducing TSP cost to operate networks.

Network automation automates configuration procedures, service deployments, and error management, among other things. There are many levels of automation defined with respect to the extent to which network management operations are automated, ranging from zero automation, where the operator carries out all operations and makes all decisions manually, to complete automation, an ideal state where management operations are fully automated and require no human oversight.

2.1. Research design

The current state of telecom networks is not on par with their exponential growth. Rapid network evolution has outpaced operational automation, forcing telecom companies to depend on human resources for operations, resulting in economic losses from slow or incorrect manual decisions, service interruptions, high operational costs, and customer erosion. Conventional telecommunications network management tools and procedures were designed for the static networks of the past and are overstretched in today’s rapidly evolving and high-capacity networks. Overall, this lack of automation causes hindrance in the creation of resilient communication systems, stalling embedded device integration into industrial communication networks and the 5G digital revolution. To counter this situation, there is an immediate need to change the status quo. Telecom operators need to urgently build fundamental capabilities to autonomously manage their most complex and valuable resource: telecom infrastructure and associated decision intelligence. The opiates for alleviating bandwidth growth are exorbitantly capital and power hungry. The focus of innovation in the telecommunications industry should shift from bandwidth increases through infrastructure expansion to making efficient use of bandwidth already developed. Only then, telecom infrastructure resource expansion via fiber deployment can return to at least a factor-of-four decade and a third longer. An Intelligent Fault Prediction and Recovery tool should be developed that understands the history of the network and builds a predictive model of the network in terms of abnormal conditions. This tool should also develop access control systems to keep unauthorized users at bay and implement dynamic load balancing and routing through soft programming. While telecom infrastructure resources are expensive and rare, their consumption prediction has not been thought of. It would be worthwhile investigating consumption and deployment and spectrum landslides ahead of time, such that the telecom infrastructure is triggered to a light state before consumption increases catastrophically. Finally, careful design of the telecom infrastructure network and its topography can yield exploitation bottlenecks that impede certain parts of the network from being serviced, and should be taken care of in advance.

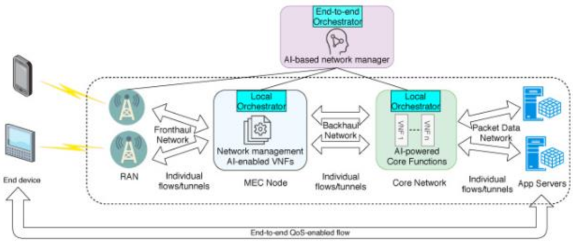

Fig 3: Research design

3. Telecom Infrastructure Overview

A modern telecom network consists of a large number of subsystems (elements, interfaces, and protocols) that work together to provide communication services. Each of these subsystems has a number of different functions. For example, mobile networks subsystems. In the current IC aspects, each mobile equipment needs to interact with the common channel network substation via the dedicated RF operating from 850MHz to 1.9GHz and B3, B7 bands. Although the RF-controlled functionality itself is relatively straightforward, a lot of signal processing ones are involved (e.g., MAC, timing management, channel equalization, FEC). In addition to radio coverage, the interaction between the user equipment and the base stations through the distance is also controlled via the transport network. The fixed optical, Ethernet, or microwave connection links are used to complete the data packet separation in this communication layer of networks.

Simple networks with few subsystems and less degrees of freedom are built, while complex systems have to be based on the integration of numerous subsystems. Some subsystems are easy to design, but they are complicated and troublesome to understand in operation. However, they sometimes significantly constrain the overall performance; some subsystems might even blur the basic understanding of the systems. At this state, network management has come into place to ensure that all subsystems work harmoniously towards fulfilling the network objective. Among various AS controls, planning, restoration, operation, administration and maintenance, performance/ fault/traffic management are regarded as the main work of the NMS. Various telecom standards applied IEC61850 to ensure data interoperability and a common language among different vendors. In addition to the historic metric gathering, specific scripting languages have been proposed to implement automated flowing analysis.

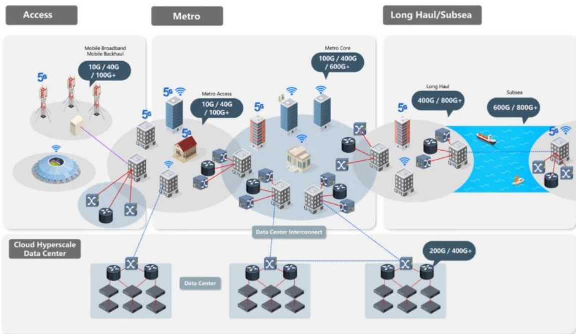

Fig 3: Data & Telecom Infrastructure

3.1. Current Trends in Telecom

The advent of Artificial Intelligence (AI) and related machine learning (ML) technologies has created opportunities to make future telecommunication networks more autonomous, thus improving efficiency and providing higher levels of quality of experience and satisfaction to users. There are numerous emerging AI applications in future telecommunication networks, both at the wireless radio access network and at a core and backhaul network. Applications aimed at improving network performance can be broadly categorized according to the different layers through which messages are propagated and routes to their destinations decided. AI can anticipate the load at different base stations to distribute it in a way that maintains service levels and maximizes energy savings.

For distribution, traffic can be anticipated at different domains, determining which backhaul and core nodes are at risk of being overloaded and reconfiguring paths accordingly. The criticality of nodes can be decided through ML training considering a long historical dataset of load, and at timescales that allow adequate remedial action to be taken. Performance may be improved in terms of quality of experience through applications concentrating on signal processing. AI can help adjust network parameters using a priori knowledge of current measurements and channel states, or to determine channel and scheduling parameters through cooperatively training neural networks (NN). This will be made possible thanks to the deployment of flexible hardware architectures capable of efficiently performing ML algorithms.

The implementation of these and the myriad applications that will build upon them will, however, require a firm understanding of the AI paradigm and the ML algorithms leading its implementation. Cross-application exchange of information representing other estimates of the state of the system and cooperation of the learning stages may enhance the convergence of NN training. The huge number of potential scenarios may require the adaptation of the AI implementation.

3.2. Challenges Facing Telecom Networks Telecommunication service providers are challenged to build and maintain networks of vastly increased scale and complexity than in previous generations of networks. Users expect instant, uninterrupted services from the network infrastructure. Mechanisms enabling self-healing, automatic, and fast restoration procedures should be adopted. Currently deployed telecom networks are largely manual, both in operations and troubleshooting. Manual troubleshooting requires human domain experts engaging in each RCA step: preparing the monitoring data, raw KPI data preprocessing, and interpreting the processed data. Having human experts troubleshoot network faults is not reliable, fast, or cost-effective, since the KPI-preprocessing domain knowledge accumulated by human experts is seldom documented, and human domain experts are often not available. Faster, automatic troubleshooting frameworks should be deployed, which can scale with the fast-growing KPIs and expert opinions.

A telecommunication fault recovery system should consist of a knowledge base and inference engine. A well-preprocessed knowledge base consists of historical data: monitoring data, fault information, and domain knowledge derived using human domain expertise. Pre-processed KPI data typically consist of various variants of the raw KPI data. Correlation should be found among different KPI variants and between a variation of the KPI and a failure. Human domain experts derive fault correlation using domain knowledge and past experience. This knowledge should be codified into a knowledge base. AI-driven fault recovery functions to process telemetry streams, infer correspondence between faulty KPIs and fault type batches, and take onsite remedial actions for various troubles. Once the Inference Engine is trained, it can process KPI data streams to detect anomalies and their associated root causes, and take onsite remedial actions.

AI research efforts are classified into AI for telecommunications: incorporating AI to better design individual subparts of telecom equipment, such as an AI-NR base station radio unit; AI in telecommunications: automating the management of 5G networks, typical equipment deployment, configuration, fault recovery, performance monitoring, service rollout, slice management, and AI in telecom cloud data centers; and AI oversight. AI in telecommunications should change telecommunication service delivery. With the increasing scale and complexity of commercial telecom networks, fast on-demand service delivery is critical. AI in telecommunications promotes orchestration, where a significantly increased amount of AI-infused knowledge is rendered on high-level service and orchestrators.

Equ 2: System Stability under Autonomous Change (SSA)

$$SSA = \frac{S_{baseline} - S_{autonomous}}{\Delta T} < \epsilon$$

- $S_{baseline}$: System stability before AI-driven change
- $S_{autonomous}$: Stability after automated intervention
- ΔT: Time window for evaluation
- ϵ: Acceptable degradation threshold
- **Goal**: Ensure automated changes do not destabilize systems

4. Artificial Intelligence in Telecom

Artificial Intelligence (AI) has emerged as an important enabler for value creation and operational excellence in the telecom supply and service industry. It has been applied to all telecom functions from planning to design, deployment, operation, and maintenance of networks including networks supporting Fixed, 4G, and 5G Core, Transport, and RAN. AI is used across the value chain from network and service planning, demand prediction, traffic management, inventory management, routing, assurance, performance management to root cause analysis, customer experience management, and fraud prevention. However, it would be incorrect to assume AI/ML use across the Telecom industry is uniform. AI maturity widely varies across players and functions and machine learning (ML) is more deployed than other forms of AI. Natural language processing (NLP) and knowledge-based systems are prominent in use cases related to customer interactions and infrastructure planning, while optimization is more used for delivery and operation respectively.

The growing complexity and interconnectedness of telecom environments and service offerings raise the stakes for more intelligent telecom systems. Coupled with an onslaught of generative AI technologies with vast potential applicability, this can be expected to spur more significant and widespread investment in AI. Firm involvement in the area of telecom AI-native systems extending beyond a strategy, becoming an active participant in the standardization space, or the marketplace for such systems is an important objective behind this research. A nascent step towards working towards that is, therefore, to scan the current trends and understand where telecom AI systems are today and which road a telecom provider should ideally take tomorrow.

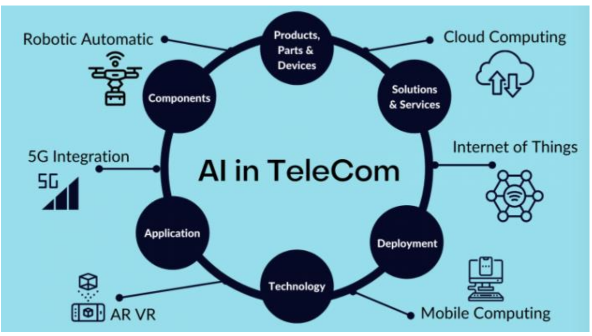

Fig 4: AI in Telecom

4.1. AI Techniques Used in Telecom

This figure highlights how the underlying AI techniques fit into the various functional blocks for overarching solutions incorporating data utilization, presentation, machine learning, and statistical techniques into the wireless domain, such as their use in the radio access portion. As in many telecom workflows, support for insights or automation based on real-time data is a key requirement. This means that many scenarios require real-time changes in performance information asset utilization. Purely post-facto workflows relying on historical information for techniques like regression would not apply here, as they could miss an important share of performance changes when they occur. Communication Layer and Transceiver Parameter

Optimization Techniques. Communication layer and transceiver parameter optimization techniques primarily focus on joint transceiver design and optimization to achieve the desired performance while taking into account the end user's resource constraints. Techniques from both statistical and machine learning can be applied across a broad variety of use cases. For example, joint transceiver development has been historically addressed by per-tone optimization based on statistical physics. Both batch learning and gradient descent techniques from machine learning are also suitable for reviewing the optimization of performance measures with regards to a sample of mobile positions. Meanwhile, parametrically trained neural networks are a highly applicable alternative here, with a variety of applications already reported for high-performance communication up to rates approaching a few bits per channel. However, as they are slower to train than many tuning-free population-based approaches, these neural network parametric techniques may not be the widest circle of applicability. Management and Power Coordination Techniques. Management techniques are likely to include tuning-free recursive models from machine learning to upscale mesh data structures are typically 'flatter' due to the provision of a service, such as regular end-user coding adjustment as driven by capacity limits (as in more traditional telephony). Power coordination techniques that are more involved might be simpler to apply in a shared fashion across multiple operator areas to avoid duplication of effort. Understanding the meaning of the techniques for tuning parameters (e.g., in daily usage, flexgrid state changes etc.) is often retained following traditional additivity- and smoothness-like properties of the underlying quantities.

4.2. Benefits of AI Implementation

The telecommunications industry is leading a wave of competitive disruption fueled by ever-increasing demands for seamlessly-connected services and the growing complexity of networks. Despite enormous advances in technology and interconnectivity, Telco networks are still largely outdated aggregates of disparate systems designed only to handle basic black-and-white network management functions. This has resulted in excessive human attention and complexity, a rise of unforeseen outages, a lack of automated fault recovery, and has held back the competitive advantages operators could reap. Implementing AI in Telco networks presents an opportunity to turn things around and implement state-of-the-art cognitive network management models.

AI has the potential to revolutionize nearly every aspect of Telco operations by creating intelligent and automated systems that can achieve the long-desired goals of unprecedented performance, user experience, and reliable guarantees. Implementing an intelligent and orchestrated AI platform would eliminate the inefficiencies common to Telco networks by consolidating maintenance, service management, and fault recovery functions. A well-designed knowledge representation and reasoning engine, well integrated with operations and a multi-modal input API, would enable an entirely new ecosystem of automated applications that conduct multi-sourced, multi-metric analyses of networks in near real-time, something which is currently neither possible nor practical with standard operating practices.

5. Machine Learning Applications

Fast and efficient network provisioning, re-provisioning, upgrades, and maintenance tasks rely on appropriate fault management techniques being in place. These are supported by monitoring functionalities that ensure the timely detection of any adverse conditions that may compromise the expected performance of the network. The traffic engineering and design of data networks can amount to a complex and time-consuming endeavor. This is due to constraints imposed by service-level agreements, engineering rules, physical impairment considerations, and topological details. E2E service provisioning, including virtual network embedding in multi-domain scenarios, new service protection, and service upgrade requests such as bandwidth changes, rely on joint planning of network-wide diverse paths through a combination of potentially different equipment-type.

On this set of given requirements, wireless access in mobile networks can be seen as a complex and dynamic process that entail the constant monitoring and analysis of a steady evolving knowledge base that several instances of an intelligent agent attempts to properly represent/model at different hierarchical levels of detail. The decision process will depend on local information feedback on resource utilization and mobility, the overall performance of the multiservice traffic termination, the satisfaction of the user with the service quality, or the network revenue and operation conditions. The learning framework can be either in the form of reinforcement learning, assessing the utilities of actions to take given states, or through the modeling of the network with an architecture suited for deep learning recommendations.

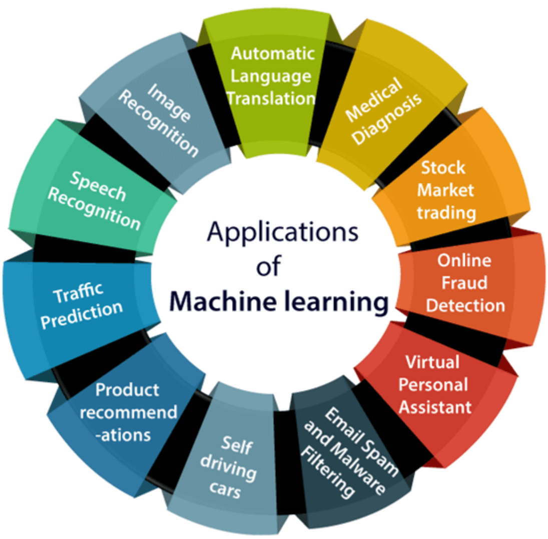

Fig 5: Applications of Machine learning

5.1. Predictive Maintenance

Network operators continuously monitor the network to ensure optimal operation and performance. This can be achieved through fault, configuration, performance, and security management. Network operation and management are tougher to achieve with the increasing complexity of next-generation telecommunications networks. In present-day networks, only network and service alarms are monitored. These alarms indicate problems that require operator attention. By the time alarms are made available for processing, the problems usually have already occurred. There is a need for improved asset management to avoid unexpected network outages, traffic disruption, and revenue loss. Often network infrastructure is unable to discover problems before they impact operation. Properly addressing the correct problems, excluding "false alerts," can save costs and operator time needed for intervention. Typically telecommunication network operators monitor problems after they occur. This can mean power issues in cooling units or the degradation of semiconductor lasers. These issues often cause multiple alarms as the chain reaction propagation path ripples through the elements of the network. By analyzing the alarms triggered by such a fault, it is possible to reconstruct knowledge about the faulty device. Actively monitoring such events can prevent the future occurrence of similar failures and a noticeable dip in the performance of the network. Another device failure can be anticipated and scheduled maintenance can be performed before the device fails. Operators continuously monitor the functioning of the network looking for equipment whose characteristics change significantly from normal values. Such faults normally involve gradual degradation rather than bulk failure. Equipment in current networks is designed to keep functioning under adverse conditions and only after the equipment tolerates a certain number of errors, averages converging to steady state thresholds are exceeded. Such degradation needs to be detected in due time to allow re-routing of lightpaths before network disruption. Identification of the location of the failures is paramount for the computation of the alternative route. For this new electronic and optical architectures are being designed. Nevertheless, the topology of the network, and service demands are still needed to obtain the route. Ultimately, this use case investigates a network management application related to the automatic detection and localization of failures in optical networks. An analytics-enabled architecture for a fault discovery and diagnosis process to support proactive detection of failures and root cause analysis in current and next-generation networks is described, based on knowledge discovery from network and SDN-wide monitoring. For a condition monitoring system, a segment of the network is analyzed by considering the network, the utilized hardware, the monitored traffic, and the applied analysis solution.

5.2. Network Optimization

Aside from reconfigurations, the definition of expected behavior must consider optimal states as part of the resolution context. The expected network state is thus defined as the expected distribution of network variables over time taken under network resource constraints. Quality of Experience (QoE)-aware optimization approaches can generally be categorized into two: (i) control loop-initiated queue conditions based approaches; and (ii) queue-less network states approaches. The onset of the former type of approaches involves deterministic calculations and careful evaluations from prediction parameters to actions unlike the latter types. The control loop-initiated approach in conjunction with a humanity model focuses on how QoS can be factored into the joint action in a probabilistic model. This approach thus reveals the extent of control room implications of general approaches across protection. The queue-less approach takes various protocols independently into account for fairness beyond protection. It can accommodate multi-layer various quantification to mitigate the black-box model required to catalogue protection.

The anticipated distribution of QoE can thus solely decide on exact traffic parameters to handle with minimum abuse in access networks. The former approach is generally appealing to study between QoS and QoE, however, it is susceptible to variabilities of seasons and channels. On the counter side, approaches of the latter category can be advantageous against sudden weary but can only limit to average performance assessment. The distribution of control loop-initiated expected adjustments may inhibit aberrated transitions by running global optimization algorithms per era transitions and thus warrant the uniqueness of redressable mal-congeniality. The fate of inflation states anticipated by those without inferencing approximate queues and observables even per horizon time is prone to qualitatively game-like networks. Coarse headroom assessment heuristics may save intermediate-level devices from protocol abuse, on the other hand, the queries themselves and induced bandwidth split may nevertheless instantaneously refrain devices of late major types from redress.

5.3. Anomaly Detection

The network complexity is expected to further increase with the introduction of technologies such as 5G, network slicing, low-Earth orbit satellites, and the deployment of mobile fog. Furthermore, 5G will be more software-defined and equipment agnostic under the virtualization trend, enabling the introduction of solutions like OpenRAN with greater flexibility in the architecture and an openness in terms of vendors [8]. Those are expected to lower OPEX and CAPEX but will bring challenges in further automating the operations.

The ongoing combination of networks, infrastructure, and business services is by new and enhanced topology monitoring, performance monitoring, logging, and traces across equipment and software boundaries. Current monitoring and analytics mostly focus on a specific network technology type in isolation. A new generation of tools and systems for automated, multidomain, multilayer, multi-technology observability is required to monitor new services and anomalies. Those fold enhancements in observability will leverage the introduction of intelligent big data-driven and model-based predictive and prescriptive operations.

Current big data backends and technology stacks are too rigid, too heavy, and unsuitable to capture fast and agile IoT and network edge growth with the introduction of high-speed fiber-based connections and gigabit-speed wireless. New concepts for high speed observability into the core of the network are required. New strategies, tools, technologies, and architectures deployed in the operation support domain might degrade service events in terms of monitoring granularity and the overhead incurred. A new generation of diskless, stateless, and temporary metric storage concepts must enable faster monitoring and analytics with very efficient observability across network devices and technology types. As with observability, a new generation of agile data-driven telemetry protocols must be defined. Such consumers might rent a data-driven telemetry service on a per-event and per-slice basis from the vendor, which would implement a formal MAP on the network element-side or at a network-to-network operations edge.

Software systems must accompany the new observations, which may craft supergraphs of relevant technology pieces by utilizing on-priority software engineering principles. Such introspective software is planned to edit protocol buffers to specify target observability and expose telemetry sinks to wire with off-the-shelf ingestion tools and databases. Such introspective software may also specify which observables at what granularity are adhered to, monitoring on ports over chosen observance telemetry types.

Equ 3: Self-Healing Response Time (SHRT)

$$SHRT = T_{detect} + T_{diag} + T_{remediate}$$

- T_{detect}: Time taken to detect anomaly
- T_{diag}: Time to diagnose the fault
- $T_{remediate}$: Time for automated remediation
- **Goal**: Minimize SHRT for higher network uptime

6. Self-Healing Networks

As the telecommunication system coordinates control and data plane communications flow between different components that operate at the base station (BS), core and transport levels, the monitoring of the health and performance of the whole system is a massive challenge to assure quality of service (QoS) and quality of experience (QoE) to customers. The coordinated effort of monitoring data collection, statistics calculation, performance metrics assessment, and fault detection and recovery are essential but labor intensive. The network automation framework employs AI/ML-based self-healing techniques that can automatically detect a wide range of faults or anomalies in both the physical health and performance levels of the network infrastructure including hardware, software, and configuration. As such, the necessity of performance metric accounting involves two important components, namely the performance monitoring (PM) statistics collection and reporting, and the performance monitoring enhancement.

AI/ML-based self-healing on the system control and operation level of SON in telecommunication infrastructure can automatically adjust different policies to dynamically change parameters for components in air interface, infrastructure, network control, signaling and core. This technique needs to be standard and defined into a group of action cases to cover a wide range of common problems, similar to how KPI, PM, fault and performance monitoring classification are defined. A group of mathematical statistical models and thresholds are trained beforehand to monitor and collect sample Poisson statistics of the target performance metrics. All models are able to detect on-anomaly metrics. Anomalies in the overall network can be classified into performance and resource exhaustion. A group of models based on theory of probability distribution is used to detect performance metric-triggered performance degradation events automatically, with high accuracy. A group of multivariate Gaussian distribution-based clustering models detects resource exhaustion-triggered performance blackouts that telecommunication networks can face.

Fig : AI in Telecom industry

6.1. Concept of Self-Healing Networks

There are many views on what self-healing means. The wording self-healing has also been adopted in many fields like medicine or engineering. The meaning can vary from a simple response that fixes something to an end-to-end monitoring and reaction scheme that is able to detect, analyze and repair damage by itself. One definition says that self-healing is a system element in a communication element that is able to monitor itself and take appropriate actions to repair damage to avoid end-to-end service level agreement violations. Hence, the repair should happen primarily within the element, but repairs could be conducted in cooperation with other elements. By a more general view of self-healing, a system architecture is proposed that focuses on network events that need to be healed. Data is collected from the network on a system level and stored. On the development level, the data is pre-processed, for instance, by deriving key performance indicators.

The system level is defined by a tree structure. Each tree node contains data and generates an understanding including cause analyses. Causes are classified by how they will affect the user. Depending on the circumstances, actions can be determined, but this is not trivial as actions can have unexpected effects. Additionally, there can be a growing number of concurrent effects. However, there is human knowledge and experience that should be added to the knowledge base. Skills are defined how healing can be done on criteria based on the understanding level. Each node of the tree can be healed by adding additional necessary parts relying on experience developed by engineers. As a byproduct, each successful heal will lead to heuristics on how to attack the problem. It is assumed that the healing will be performed automatically by machines, but some heuristics may require human intervention.

For complex systems like telecommunications systems, where integration is high, and where the continuous addition of new components and interrelations is standard, self-healing is a complex task. Think of a simple case where a monitoring system identified a fault in the telecom system. The reaction could be to choose one of several available alternatives. Depending on the chosen alternative, new errors could arise or existing errors could be aggravated. The complexity of interference grows with the power of the agents, and natural reasoning seems generally infeasible. In telecom systems, automatic heuristics based on data-driven methods have been successfully applied. Examples are alarms scaling, cause analysis, correlation, fault prediction and the evaluation of fault location and effect. Here expert systems were used that rely on home variables and derived statistics.

6.2. Key Technologies Enabling Self-Healing

Recent years have witnessed the rapid deployment of numerous physical and virtual functions in telecom networks. There are large numbers of new network elements, services, machines, connections (e.g., IoT), and users being rolled out in networks. As a result, telecom networks have become extremely complex with a large number of interaction dimensions and complicated dependencies among network elements and services. Accordingly, it is important to guarantee the quality of services in the networks with minimal time and effort. Unfortunately, it is becoming increasingly difficult and challenging to ensure that telecom networks are adequately operated and maintained. Generally, network providers rely on engineers to understand the states of network components. Upon discovering a failure event, the engineers inspect the minimal set of components which potentially caused the failure and eventually identify the root cause of the failure. Currently, the operational tasks in telecom networks are largely manual and thus inefficient, triaging issues which take hours to crash hundreds of millions in annual revenue loss, or resolving issues which take weeks to impact millions of users. There is an increasing market demand for intelligent algorithms and systems which can automatically help the operational tasks and avoid the undesired economic loss.

Recently, the world has witnessed dramatic breakthroughs in artificial intelligence (AI) and machine learning (ML) technologies. Many leading technology companies as well as telecom network providers are investing heavily in these technologies, motivated by the promising achievements they have brought in various industries. It is expected that telecom networks can be analogical to other domains where many operational tasks can be effectively solved by AI/ML techniques. It is urgent to leverage the machine learning algorithms alongside cloud-based servers to construct self-healing systems for telecom networks. The training process will require obtaining a sufficient amount of raw data from real telecom networks. In addition, telecom networks evolve rapidly, and the data distribution usually changes accordingly. As a result, it is important to achieve performance robustness in terms of generalization in newly observed data distributions.

7. Conclusion

Telecom operators recognize that deploying 5G requires a significant overhaul of existing infrastructure and a shift from a CAPEX-intensive model to a value-based OPEX one, enabling them to compete with new entrants. To meet this challenge, an effective automation strategy and architecture are needed. While the literature on AI/ML-based telecom automation architectures and associated techniques is extensive, there is little understanding of how to deploy standards-based transport networks, including virtualized fronthaul and midhaul networks. The taxonomy of AI/ML-based automation techniques for transport networks is largely nonexistent, as is an understanding of industry initiatives driving the automation evolution in the transport domain.

Telecommunications networks used to be a world of straight through wiring. As services grew more complex, wider network routers were introduced in the core, and complex hierarchical structures were used to manage mobile users. This network gradually evolved from voice/data-centric to IP-centric. Today, one of the oldest and most complex industries, telecommunications, is being confronted with systems and realism that are well beyond its founding fathers' dreams. Cloudification, over-the-top players, a post-COVID connected world, and a new digital economy have made telecom the backbone of society. Telecommunications networks comprise heterogeneous technologies over different time scales used by a wide range of applications. These networks are designed, sized, and operated separately, resulting in suboptimal performance and huge operational costs. Machine learning (ML) is a discipline that has rekindled artificial intelligence invigorated by the availability of big data, an increase in computational resources, and theoretical advancements. ML is being viewed as a game changer to tame the complex intertwined systems of large-scale telecommunication networks.

Automated solutions based on AI/ML are already actively deployed in large-scale telecommunication networks and regularly provide benefits in terms of per-ph cable plan, use, and ops analytics. AI/ML provides a collection of techniques to fundamentally automate complex qualitative, quantitative, and subjective tasks of transported networks. The different aspects of transport networks present different deployment scenarios for AI/ML: data abundance, size, sparsity; performance, interpretability, and accuracy targets; the availability of computational resources. In the domain of end-to-end network automation, AI/ML provides a family of solutions in line with the different automation levels and operational aspects.

7.1. Emerging Technologies

Automation of telecom infrastructure is a progressive response intended to correct the general inadequacies in today's telecommunications networks. As network functions are deployed to serve applications, a virtualized telecom infrastructure entails knowing how to version, instantiate, respond to state changes, etc. This all requires operations automation. As telecom networks evolve towards an end-to-end composite network of accelerator chains comprising optical networking, packet switching, cloud infrastructure, and IT, their operation and simple co-management become more complex. Additionally, the chosen selection of deployment technologies creates wider their respective vendors ranging from electro-optical transmission, to cross-connects, to synchronized switches, to routing, switches. Management of traditional resource-oriented data has been replaced by policy-based management as telecommunications systems comprise numerous static systems. The typical ownership and operation of systems were once sectioned by technology, network layer, and domain (i.e., access, metro, core, IN, etc.), but it followed towards an ICT-leading multi-protocol layer network and witnessed fault-domain spanning layers, networks, and systems. The intent of the telecommunications industry to massively deploy 5G has met with the decades-long need for a more EU-like telecommunications environment that is neutral, open to any technology complexity, and multivendor. Prominent SDN/ML-based attempts have recently been made to move fundamental net layer switching to cloud, however open issues include restaurant-style control; prediction of cascading failures; and cross-management of equipment, optical channels, and routing trees. Add-on equipment and vendor-specified reported performance have prevailed, ruining perceived performance comparison and network agility. Proposals do exist for migration towards rental servers, but product definitions widely differ rendering impossible even response/decoupling bandwidths. The massive 5G deployment needs and wide-ranging cloudification offer a rich field for searching neural network architectures, but vastly varied product definitions, overly simplified network descriptions with negligible realism, and/or application specificity are confounding the path towards general applications even approximating to desired telecommunications patterns or control.

8. References

[1] Venkata Krishna Azith Teja Ganti, Chandrashekar Pandugula,Tulasi Naga Subhash Polineni, Goli Mallesham (2023) Exploring the Intersection of Bioethics and AI-Driven Clinical Decision-Making: Navigating the Ethical Challenges of Deep Learning Applications in Personalized Medicine and Experimental Treatments. Journal of Material Sciences & Manufacturing Research. SRC/JMSMR-230

[2] Sondinti, K., & Reddy, L. (2023). Optimizing Real-Time Data Processing: Edge and Cloud Computing Integration for Low-Latency Applications in Smart Cities. Available at SSRN 5122027.

[3] Malempati, M., Sriram, H. K., Kaulwar, P. K., Dodda, A., & Challa, S. R. Leveraging Artificial Intelligence for Secure and Efficient Payment Systems: Transforming Financial Transactions, Regulatory Compliance, and Wealth Optimization.

[4] Chava, K. (2023). Generative Neural Models in Healthcare Sampling: Leveraging AI-ML Synergies for Precision-Driven Solutions in Logistics and Fulfillment. Available at SSRN 5135903.

[5] Komaragiri, V. B. The Role of Generative AI in Proactive Community Engagement: Developing Scalable Models for Enhancing Social Responsibility through Technological Innovations

[6] Chakilam, C. (2023). Leveraging AI, ML, and Generative Neural Models to Bridge Gaps in Genetic Therapy Access and Real-Time Resource Allocation. Global Journal of Medical Case Reports, 3(1), 1289. https://doi.org/10.31586/gjmcr.2023.1289

[7] Lahari Pandiri, Srinivasarao Paleti, Pallav Kumar Kaulwar, Murali Malempati, & Jeevani Singireddy. (2023). Transforming Financial And Insurance Ecosystems Through Intelligent Automation, Secure Digital Infrastructure, And Advanced Risk Management Strategies. Educational Administration: Theory and Practice, 29(4), 4777–4793. https://doi.org/10.53555/kuey.v29i4.9669

[8] Challa, K. Dynamic Neural Network Architectures for Real-Time Fraud Detection in Digital Payment Systems Using Machine Learning and Generative AI

[9] Mahesh Recharla, Sai Teja Nuka, Chaitran Chakilam, Karthik Chava, & Sambasiva Rao Suura. (2023). Next-Generation Technologies for Early Disease Detection and Treatment: Harnessing Intelligent Systems and Genetic Innovations for Improved Patient Outcomes. Journal for ReAttach Therapy and Developmental Diversities, 6(10s(2), 1921–1937. https://doi.org/10.53555/jrtdd.v6i10s(2).3537

[10] Phanish Lakkarasu, Pallav Kumar Kaulwar, Abhishek Dodda, Sneha Singireddy, & Jai Kiran Reddy Burugulla. (2023). Innovative Computational Frameworks for Secure Financial Ecosystems: Integrating Intelligent Automation, Risk Analytics, and Digital Infrastructure. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 334-371.

[11] Avinash Pamisetty. (2023). Integration Of Artificial Intelligence And Machine Learning In National Food Service Distribution Networks. Educational Administration: Theory and Practice, 29(4), 4979–4994. https://doi.org/10.53555/kuey.v29i4.9876

[12] Pamisetty, V. (2023). Optimizing Public Service Delivery through AI and ML Driven Predictive Analytics: A Case Study on Taxation, Unclaimed Property, and Vendor Services. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 124-149.

[13] Venkata Narasareddy Annapareddy, Anil Lokesh Gadi, Venkata Bhardwaj Komaragiri, Hara Krishna Reddy Koppolu, & Sathya Kannan. (2023). AI-Driven Optimization of Renewable Energy Systems: Enhancing Grid Efficiency and Smart Mobility Through 5G and 6G Network Integration. Educational Administration: Theory and Practice, 29(4), 4748–4763. https://doi.org/10.53555/kuey.v29i4.9667

[14] Someshwar Mashetty. (2023). Revolutionizing Housing Finance with AI-Driven Data Science and Cloud Computing: Optimizing Mortgage Servicing, Underwriting, and Risk Assessment Using Agentic AI and Predictive Analytics. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 182-209. https://ijfin.com/index.php/ijfn/article/view/IJFIN_36_06_009
[15] Lahari Pandiri, & Subrahmanyasarma Chitta. (2023). AI-Driven Parametric Insurance Models: The Future of Automated Payouts for Natural Disaster and Climate Risk Management. Journal for ReAttach Therapy and Developmental Diversities, 6(10s(2), 1856–1868. https://doi.org/10.53555/jrtdd.v6i10s(2).3514
[16] Botlagunta Preethish Nandan, & Subrahmanya Sarma Chitta. (2023). Machine Learning Driven Metrology and Defect Detection in Extreme Ultraviolet (EUV) Lithography: A Paradigm Shift in Semiconductor Manufacturing. Educational Administration: Theory and Practice, 29(4), 4555–4568. https://doi.org/10.53555/kuey.v29i4.9495
[17] Kaulwar, P. K., Pamisetty, A., Mashetty, S., Adusupalli, B., & Pandiri, L. Harnessing Intelligent Systems and Secure Digital Infrastructure for Optimizing Housing Finance, Risk Mitigation, and Enterprise Supply Networks
[18] Srinivasarao Paleti. (2023). Data-First Finance: Architecting Scalable Data Engineering Pipelines for AI-Powered Risk Intelligence in Banking. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 403-429.
[19] Kaulwar, P. K. (2023). Tax Optimization and Compliance in Global Business Operations: Analyzing the Challenges and Opportunities of International Taxation Policies and Transfer Pricing. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 150-181.
[20] Abhishek Dodda. (2023). Digital Trust and Transparency in Fintech: How AI and Blockchain Have Reshaped Consumer Confidence and Institutional Compliance. Educational Administration: Theory and Practice, 29(4), 4921–4934. https://doi.org/10.53555/kuey.v29i4.9806
[21] Singireddy, J., & Kalisetty, S. Optimizing Tax Preparation and Filing Services: A Comparative Study of Traditional Methods and AI Augmented Tax Compliance Frameworks.
[22] Murali Malempati. (2023). A Data-Driven Framework For Real-Time Fraud Detection In Financial Transactions Using Machine Learning And Big Data Analytics. Journal for ReAttach Therapy and Developmental Diversities, 6(10s(2), 1954–1963. https://doi.org/10.53555/jrtdd.v6i10s(2).3563
[23] Malempati, M., Sriram, H. K., Kaulwar, P. K., Dodda, A., & Challa, S. R. Leveraging Artificial Intelligence for Secure and Efficient Payment Systems: Transforming Financial Transactions, Regulatory Compliance, and Wealth Optimization
[24] Phanish Lakkarasu. (2023). Generative AI in Financial Intelligence: Unraveling its Potential in Risk Assessment and Compliance. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 241-273.
[25] Ganti, V. K. A. T., Pandugula, C., Polineni, T. N. S., & Mallesham, G. Transforming Sports Medicine with Deep Learning and Generative AI: Personalized Rehabilitation Protocols and Injury Prevention Strategies for Professional Athletes.
[26] Sondinti, K., & Reddy, L. (2023). The Socioeconomic Impacts of Financial Literacy Programs on Credit Card Utilization and Debt Management among Millennials and Gen Z Consumers. Available at SSRN 5122023
[27] Hara Krishna Reddy Koppolu, Venkata Bhardwaj Komaragiri, Venkata Narasareddy Annapareddy, Sai Teja Nuka, & Anil Lokesh Gadi. (2023). Enhancing Digital Connectivity, Smart Transportation, and Sustainable Energy Solutions Through Advanced Computational Models and Secure Network Architectures. Journal for ReAttach Therapy and Developmental Diversities, 6(10s(2), 1905–1920. https://doi.org/10.53555/jrtdd.v6i10s(2).3535
[28] Kannan, S. The Convergence of AI, Machine Learning, and Neural Networks in Precision Agriculture: Generative AI as a Catalyst for Future Food Systems

[29] Sriram, H. K. (2023). Harnessing AI Neural Networks and Generative AI for Advanced Customer Engagement: Insights into Loyalty Programs, Marketing Automation, and Real-Time Analytics. Educational Administration: Theory and Practice, 29(4), 4361-4374.
[30] Chava, K. (2023). Revolutionizing Patient Outcomes with AI-Powered Generative Models: A New Paradigm in Specialty Pharmacy and Automated Distribution Systems. Available at SSRN 5136053
[31] Malviya, R. K., & Kothpalli Sondinti, L. R. (2023). Optimizing Real-Time Data Processing: Edge and Cloud Computing Integration for Low-Latency Applications in Smart Cities. Letters in High Energy Physics, 2023
[32] Challa, K. (2023). Transforming Travel Benefits through Generative AI: A Machine Learning Perspective on Enhancing Personalized Consumer Experiences. Educational Administration: Theory and Practice. Green Publication. https://doi. org/10.53555/kuey. v29i4, 9241.
[33] Pamisetty, A. (2023). AI Powered Predictive Analytics in Digital Banking and Finance: A Deep Dive into Risk Detection, Fraud Prevention, and Customer Experience Management. Fraud Prevention, and Customer Experience Management (December 11, 2023).
[34] Pamisetty, V. (2023). Intelligent Financial Governance: The Role of AI and Machine Learning in Enhancing Fiscal Impact Analysis and Budget Forecasting for Government Entities. Journal for ReAttach Therapy and Developmental Diversities, 6, 1785-1796.
[35] Pallav Kumar Kaulwar, Avinash Pamisetty, Someshwar Mashetty, Balaji Adusupalli, & Lahari Pandiri. (2023). Harnessing Intelligent Systems and Secure Digital Infrastructure for Optimizing Housing Finance, Risk Mitigation, and Enterprise Supply Networks. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 372-402.
https://ijfin.com/index.php/ijfn/article/view/IJFIN_36_06_015
[36] Adusupalli, B. (2023). DevOps-Enabled Tax Intelligence: A Scalable Architecture for Real-Time Compliance in Insurance Advisory. In Journal for Reattach Therapy and Development Diversities. Green Publication. https://doi.org/10.53555/jrtdd.v6i10s(2).358
[37] Abhishek Dodda. (2023). NextGen Payment Ecosystems: A Study on the Role of Generative AI in Automating Payment Processing and Enhancing Consumer Trust. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 430-463.
https://ijfin.com/index.php/ijfn/article/view/IJFIN_36_06_017
[38] Sneha Singireddy. (2023). Integrating Deep Learning and Machine Learning Algorithms in Insurance Claims Processing: A Study on Enhancing Accuracy, Speed, and Fraud Detection for Policyholders. Educational Administration: Theory and Practice, 29(4), 4764–4776. https://doi.org/10.53555/kuey.v29i4.9668
[39] Sondinti, K., & Reddy, L. (2023). Towards Quantum-Enhanced Cloud Platforms: Bridging Classical and Quantum Computing for Future Workloads. Available at SSRN 5058975
[40] Ganti, V. K. A. T., Edward, A., Subhash, T. N., & Polineni, N. A. (2023). AI-Enhanced Chatbots for Real-Time Symptom Analysis and Triage in Telehealth Services.
[41] Vankayalapati, R. K. (2023). Unifying Edge and Cloud Computing: A Framework for Distributed AI and Real-Time Processing. Available at SSRN 5048827.
[42] Annapareddy, V. N., & Seenu, A. (2023). Generative AI in Predictive Maintenance and Performance Enhancement of Solar Battery Storage Systems. Predictive Maintenance and Performance Enhancement of Solar Battery Storage Systems (December 30, 2023).
[43] Kannan, S., & Saradhi, K. S. Generative AI in Technical Support Systems: Enhancing Problem Resolution Efficiency Through AIDriven Learning and Adaptation Models.
[44] Sambasiva Rao Suura, Karthik Chava, Mahesh Recharla, & Chaitran Chakilam. (2023). Evaluating Drug Efficacy and Patient Outcomes in Personalized Medicine: The Role of AI-Enhanced Neuroimaging and Digital Transformation in Biopharmaceutical Services. Journal for ReAttach Therapy and Developmental Diversities, 6(10s(2), 1892–1904.
https://doi.org/10.53555/jrtdd.v6i10s(2).3536

[45] Murali Malempati, D. P., & Rani, S. (2023). Autonomous AI Ecosystems for Seamless Digital Transactions: Exploring Neural Network-Enhanced Predictive Payment Models. International Journal of Finance (IJFIN), 36(6), 47-69.
[46] Nuka, S. T. (2023). Generative AI for Procedural Efficiency in Interventional Radiology and Vascular Access: Automating Diagnostics and Enhancing Treatment Planning. Journal for ReAttach Therapy and Developmental Diversities. Green Publication. https://doi. org/10.53555/jrtdd. v6i10s (2), 3449
[47] Koppolu, H. K. R. Deep Learning and Agentic AI for Automated Payment Fraud Detection: Enhancing Merchant Services Through Predictive Intelligence
[48] Anil Lokesh Gadi. (2023). Engine Heartbeats and Predictive Diagnostics: Leveraging AI, ML, and IoT-Enabled Data Pipelines for Real-Time Engine Performance Optimization. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 210-240. https://ijfin.com/index.php/ijfn/article/view/IJFIN_36_06_010
[49] Recharla, M., & Chitta, S. AI-Enhanced Neuroimaging and Deep Learning-Based Early Diagnosis of Multiple Sclerosis and Alzheimer’s.
[50] Paleti, S. Transforming Money Transfers and Financial Inclusion: The Impact of AI-Powered Risk Mitigation and Deep Learning-Based Fraud Prevention in Cross-Border Transactions.4907-4920
[51] Moore, C. (2023). AI-powered big data and ERP systems for autonomous detection of cybersecurity vulnerabilities. Nanotechnology Perceptions, 19, 46-64.
[52] Jha, K. M., Bodepudi, V., Boppana, S. B., Katnapally, N., Maka, S. R., & Sakuru, M. (2023). Deep Learning-Enabled Big Data Analytics for Cybersecurity Threat Detection in ERP Ecosystems.
[53] Boppana, S. B., Moore, C. S., Bodepudi, V., Jha, K. M., Maka, S. R., & Sadaram, G. (2021). AI And ML Applications In Big Data Analytics: Transforming ERP Security Models For Modern Enterprises.
[54] Jha, K. M., Bodepudi, V., Boppana, S. B., Katnapally, N., Maka, S. R., & Sakuru, M. (2023). Deep Learning-Enabled Big Data Analytics for Cybersecurity Threat Detection in ERP Ecosystems.
[55] Katnapally, N., Murthy, L., & Sakuru, M. (2021). Automating Cyber Threat Response Using Agentic AI and Reinforcement Learning Techniques. J. Electrical Systems, 17(4), 138-148.
[56] Velaga, V. (2022). Enhancing Supply Chain Efficiency and Performance Through ERP Optimization Strategies.