RESEARCH PAPER

Smart Enterprise Cloud Computing Using Artificial Intelligence and Machine Learning Techniques: An Analytical Study of Intelligent Resource Optimization, Predictive Security, and Business Automation

Apoorv Singh[1*], Dr. Bharat Bhushan Agarwal[2]

[1*]*Research Scholar, Department of CSE, IFTM University, Moradabad, India.*
Email: apoorvsingh1001@gmail.com

[2]*Professor, Department of CSE, IFTM University, Moradabad, India.*
Email: bharatagarwal9@gmail.com

ABSTRACT

Enterprise cloud computing has advanced from a capacity replacement model to a smart operating model where telemetry, automation, security analytics and business workflows are integrated continuously. The paper is mainly focused on the role that AI and ML based solutions have on intelligent enterprise cloud computing through techniques such as workload forecasting, adaptive resource provisioning, security predictions and business automation. The researcher uses a mixed analytical and secondary-data-based research design. The article synthesizes peer-reviewed literature, cloud architecture frameworks, security standards and public workload/security datasets including Google Borg traces, Alibaba cluster traces, Microsoft Azure public traces and CSE-CIC-IDS2018 on AWS. As we did not download and empirically model any raw traces in this manuscript, the paper presents a verified dataset-mapping and analysis framework as opposed to fictitious performance values. According to the analysis, regression models, decision trees, random forests, neural networks, reinforcement learning, clustering, isolation forests, autoencoders and time-series models can help achieve different cloud-management tasks. Nevertheless, the actual benefit will depend on data quality, observability coverage, governance maturity and explainability. The model under consideration combines telemetry collection with Artificial Intelligence and Machine Learning analytics along with intelligent orchestration. The intention of the paper is to contribute an IEEE-style conceptual and analytical framework that connects AI-enabled resource optimization, ZTO oriented predictive security and business automation. Further, it emphasizes verifiable data, cautious interpretation and empirical validation. Future research should evaluate the framework with production-grade workload traces, sector-specific cloud cost data and amenable security-event datasets.

Keywords: Smart Enterprise, Cloud Computing, Artificial Intelligence, Machine Learning, Resource Optimization, Predictive Security, Business Automation, AIOps.

How to cite this article: Singh A, Agarwal BB. Smart Enterprise Cloud Computing Using Artificial Intelligence and Machine Learning Techniques: An Analytical Study of Intelligent Resource Optimization, Predictive Security, and Business Automation. Int J Drug Deliv Technol. 2026;16(57s): 1732-1744. DOI: 10.25258/ijddt.16.57s.174

1. INTRODUCTION

Cloud computing has emerged as a cornerstone model for enterprise information systems. We can therefore define cloud computing as an on-demand network access to a pool of configurable computing resources that can be provisioned and released with minimal management effort [1]. For businesses, moving from owned data centers to cloud-native environments has altered infrastructure planning from static capacity procurement to elastic service consumption. Nevertheless, merely being elastic, an enterprise cloud is not intelligent. To be a smart enterprise cloud, the cloud must sense changes in workloads and learn from operational patterns, automate provisioning decisions, detect adverse behavior in security, and translate the technical metric into business metrics.

The modern-day enterprise cloud is diverse too. The workload is spread among various platforms including IaaS, PaaS, SaaS, container orchestration platforms, serverless services and edge-cloud. The operational impact is profound: telemetry volume expands, failure modes distribute, identities span multiple control planes, and cost control requires ongoing attribution. Frameworks for public cloud architecture emphasize operational excellence, security, reliability, performance efficiency, cost optimization, and sustainability as core concerns for workload design [2]–[4]. It is impossible to effectively tackle these issues with manual oversight alone, as cloud occurrences usually take place at machine speed.

AI and machine learning have increasingly been installed as control layers on clouds. Machine learning can help manage resources in demand forecasting and support autoscaling support through an early warning process. In security, automated analytics can through anomalies detect suspicious internal or external behavior. In business process automation, AI-enabled analytics can be used to connect cloud operations to workflow execution, service-level management and managerial decision support. The challenge is not whether it is possible to use AI in the cloud, but how to embed AI into enterprise cloud governance without creating opaque, costly or unsafe automation. The conceptual and analytical framework based on authentic literature, official frameworks and public datasets can assist

adequately in addressing the problem.

The paper is deliberately cautious with regard to numerical claims. It does not provide false values on accuracy, cost savings, or workload. This work does not propose its own analytical procedure; rather, it identifies valid public databases and lays out an analytical design in which empirical results can be inserted upon validation of the datasets. This is important for IEEE-style research integrity as the results of cloud performance and security are considerably impacted by the workload type, data preprocessing, feature definition, baseline selection, drawback deployment context.

2. RESEARCH BACKGROUND

The service models, deployment models and operational paradigms explain enterprise cloud computing. According to the National Institute of Standards and Technology (NIST) software-as-a-service, platform-as-a-service and infrastructure-as-a-service are service models and the cloud deployment models include public, private, community and hybrid clouds [1]. In business use , these models have grown into multicloud, edge-cloud and serverless. The adoption of multicloud helps in the distribution of workloads to multiple providers, while edge-cloud systems bring computation closer to the end user or devices. Serverless computing abstracts capacity management and elevates the significance of event-driven observability.

The development of cloud platforms also has resulted in new operations disciplines. The AWS Well-Architected Framework delineates the design of the cloud in the six pillars of operational excellence, security, reliability, performance efficiency, cost optimization and sustainability [2]. The Azure Well-Architected Framework likewise focuses on the same key aspects as above [3]. The Google Cloud Well-Architected Framework includes operational excellence, security/privacy/compliance, reliability, cost optimization and performance optimization [4]. The Business Case Dynamic Framework and the Cloud Computing Design Framework demonstrate that smart enterprise cloud design is not only of compute allocation, but also a general multi-objective governance problem.

The implementation of AIOps, predictive analytics, adaptive scheduling, security analytics and automation helps in entry of AI and ML into the cloud environment. AIOps merges artificial intelligence (AI) and operational data from cloud infrastructures for incident detection, failure prediction, root-cause analysis and automated action [16]. Public trace datasets from Google, Alibaba and Microsoft Azure show that large-scale cloud systems generate rich telemetry related to jobs, tasks, virtual machines, serverless invocations, resource usage and scheduling decisions. [9]-[11] Datasets utilized in intrusion detection include CSE-CIC-IDS2018 on AWS which contains labelled network and log features useful for ML [12]. These datasets make it possible to empirically test smart-cloud models if they are correctly cleansed, interpreted and validated.

Integration of infrastructure, analytical and governance assessments needed for enterprise smart cloud systems. Although workload forecasting may determine scaling actions, a governance layer must validate the cost, compliance and SLA impacts. Anomaly models that flag strange flows must be coupled with a security layer that connects those signals to identity, policy and response playbooks. Business automation may speed up workflows, but active human supervision remains necessary where automated actions affect customers, regulated data or critical business operations.

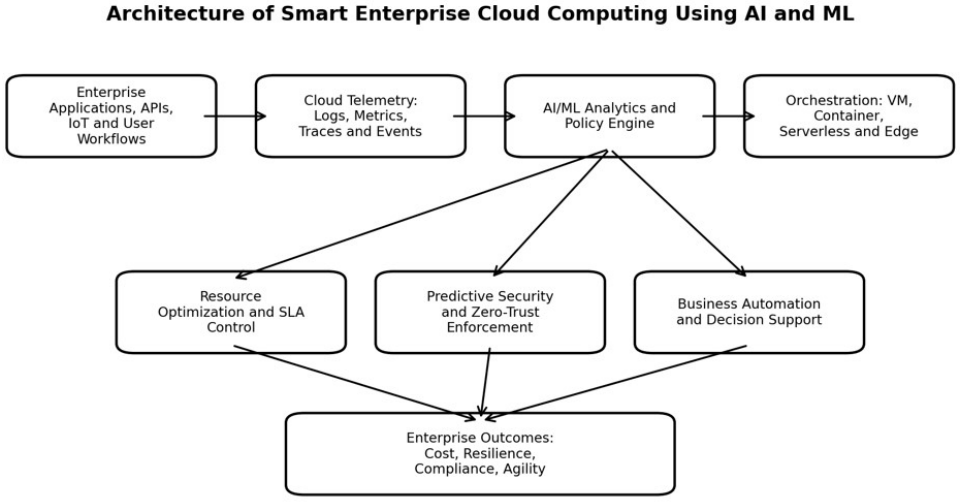

Fig. 1. Architecture of smart enterprise cloud computing using AI and ML.

TABLE 1. SUMMARY OF MAJOR CLOUD COMPUTING MODELS AND ENTERPRISE USE CASES

Model	Core Meaning	Enterprise Use Case	Smart-Cloud Relevance

IaaS	Virtualized compute, storage and networking	Migration of enterprise workloads and custom applications	AI can optimize VM sizing, placement and scaling [1].
PaaS	Managed runtime, database and development services	Faster application delivery and API development	ML can forecast platform capacity and detect service anomalies.
SaaS	Provider-managed application delivered over network	CRM, ERP, HRM, collaboration and analytics	AI can automate workflows and user-support functions.
Hybrid cloud	Combination of private/on-premises and public cloud resources	Regulated workloads with cloud bursting	Policy analytics can balance compliance and elasticity.
Multicloud	Use of two or more cloud providers	Resilience, bargaining power and best-of-breed services	Normalized telemetry and governance reduce fragmentation.
Edge-cloud	Compute close to users/devices integrated with central cloud	IoT, low-latency analytics and industrial systems	AI supports local inference and central model governance.
Serverless	Event-driven execution without direct server management	APIs, event pipelines and bursty automation	ML can forecast invocation patterns and cold-start risk [14].

3. PROBLEM STATEMENT

With the expansion of cloud adoption, enterprises face challenges this time.

Application and Providers for Cross-Business Units Resource underutilization is having more resource capacity than actual demand in compute, memory, storage or network. Cost overruns happen when elastic consumption is unlinked to workload forecasts, tagging discipline, budget alerts or FinOps governance. Workload prediction problems lead to over-provisioning, which causes SLA violations. The rate of security threats is likely o be increased at a rate that will be unimaginable.

These issues become more pronounced in hybrid and multicloud settings. Telemetry may be split across a number of providers, policies of data governance may vary across jurisdictions, and vendor-specific toolsets lock-in users. Threshold-based rules and manual monitoring are not adequate for dynamic workloads. The overall impact of AI on the global economy could be significant, reaching as much as US $16 trillion.

Accordingly, the research problem is to create an enterprise smart cloud model that employs artificial intelligence and machine learning for optimal, secured and automatic operation, while ensuring governance, explainability and empirical accountability.

4. GOALS OF RESEARCH

Analyzing the role of AI and ML in smart enterprise cloud computing is the first objective. The second objective also highlights the use of AI to help in the optimization of cloud resources. The third objective focuses on investigating predictive security and anomaly detection in cloud systems based on machine learning. The fourth objective evaluates the contribution of smart cloud automation to business efficiency. The fifth goal seeks to understand the technical, governance and ethical challenges in AI-enabled enterprise cloud adoption. The sixth aim is to design the conceptual and analytical framework for smart enterprise cloud implementation.

5. RESEARCH QUESTIONS

What influence do AI and ML techniques have on conventional enterprise cloud computing to develop smart-cloud? RQ2: What are the most relevant AI/ML techniques for resource optimization, workload prediction and autoscaling? How ML-based anomaly detection can be used in predictive cloud security for zero trust?

Intelligent automation contributes to increasing overall business efficiency and enhancing decision-making. What challenges do privacy, explainability, vendor lock-in, compliance and skill gaps pose to adoption? RQ6: What conceptual model can integrate AI/ML analytics with enterprise cloud governance, FinOps, security and business automation?

6. REVIEW OF RELATED LITERATURE

6.1 Cloud Computing Architecture and Production-Scale Workloads

The literature on cloud computing started with virtualization, service abstraction and on-demand provisioning. Further, the need for enterprises the literature began to evolve towards architecture referring to cloud models such as PaaS, SaaS, and IaaS. NIST's definition is still important because it identifies essential characteristics including on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service[1]. Cloud research began to examine production-scale workloads. The complexity of scheduling production clusters of substantial size has been brought to the fore by an examination of Google's Borg, which shows that cluster management is greatly affected by workload arrival rates, resource overcommitment, job dependencies and heavy-tailed resource consumption [13]. According to further serverless studies, invoking functions may have implication of cold starts and highly variable patterns [14].

6.2 Workload Forecasting and Resource Management

As the performance and cost of cloud systems largely

depend on the degree of matching between provisioned and actual resources, workload forecasting and resource management have become topical issues. Surveys on workload forecasting show that the ML models can predict the future workload and assist in proactive resource allocation [17]. More recent work in the area of machine-learning-based autoscaling argues that selection of application-specific metrics can improve the behaviour of the autoscaler. However, these improvements depend on how the evaluation is designed and the context of workload used to measure the performance improvements. Deep neural methods have also been explored for multivariate workload prediction in a cloud environment [19].

6.3 Predictive Analytics in Cloud Infrastructure

Using historical telemetry, predictive analytics in cloud infrastructure anticipate future states. Cloud infrastructure is designed to store process and manage big data. Regression models can ascertain the demand from the observed variables, the time-series models can predict the workload arrivals, reinforcement learning may learn the scaling or placement policies, clustering can identify the workload classes and deep learning can model the nonlinear interactions among resources. When telemetry is reliable, these methods are useful and when validation distinguishes predictive performance from operational utility. A CPU demand prediction model may fail to reduce cost if it does not incorporate provisioning lead time and pricing model and SLA constraints.

6.4 Machine Learning for Cloud Security and Anomaly Detection

Research in cloud security has been increasingly using Machine Learning for intrusion detection, anomaly detection, and security-event prioritization. An AWS repository containing brute-force, Heartbleed, botnet, DoS, DDoS, web attacks and infiltration scenarios, network traffic, system logs and 80 CICFlowMeter features [12]. A study of feature selection on CSE-CIC-IDS2018 has shown that performing feature selection is more important than assuming all traffic attributes are equally useful [22]. In cloud computing, aside from anomaly detection, alert fatigue, benign switching, and zero-day must-behaviors must be handled in the anomaly detection solution which calls for contextual risk scoring. [23]

6.5 Cloud Automation and Intelligent Orchestration

Intelligent orchestration integrates monitoring, analysis, planning and execution. Conventional threshold-based autoscaling responds to metric breaches while AI-enabled autoscaling predicts capacity needs and acts ahead of time. Production-ready autoscaling continues to be challenging as real systems face challenges from noisy metrics, variable workload and delayed actuation. Moreover, cost-performance trade-offs and complex dependencies further complicate tasks. Hence, cloud automation should incorporate human approval paths and rollback processes for high-risk activities.

6.6 Multicloud Governance, Observability and Policy Enforcement

The article focuses on the solutions that can help in selections of observability, identity governance, cost attribution and policy enforcement in complex issues. Cloud workloads can be quite distinct as illustrated by Azure public traces, Google Borg traces and Alibaba traces in the contexts of virtual-machine, serverless GPU and batch processing [9][11]. A smart enterprise cloud must therefore avoid provider-specific assumptions We need to standardize telemetry, normalize tags, align identity policies and unify governance metrics across providers.

6.7 Cybersecurity, Zero Trust and AI Risk Governance

The NIST CSF 2.0 emphasizes the necessity of executing governance, identification, protection, detection, response and recovery as concurrent cybersecurity functions. NIST zero trust focuses on protecting resources, not segments, and not trusting implicitly based on location [6]. The management of AI-related risks to individuals, organizations and society is emphasized. Adoption of AI-HPC cloud seems not just a technical optimization problem, but also a governance issue. Importantly, we mean accountability, auditability, privacy, risk tolerance and trustworthiness.

TABLE 2. COMPARATIVE REVIEW OF EXISTING STUDIES AND SOURCES

Source	Focus	Data/Method	Relevance to This Paper
Mell and Grance [1]	Cloud definition and service/deployment models	NIST conceptual standard	Provides baseline cloud terminology.
Tirmazi et al. [13]	Google Borg production trace	May 2019 trace across eight Borg clusters	Supports workload heterogeneity and scheduling analysis.
Shahrad et al. [14]	Serverless workload characterization	Azure Functions production workload	Supports event-driven and serverless resource analysis.
Cheng et al. [15]	Alibaba co-located workloads	Production cluster trace	Shows online/batch workload interaction.
Cheng et al. [16]	AIOps on cloud platforms	Review and taxonomy	Supports incident detection, failure prediction and automated actions.
Pintye et al. [18]	ML autoscaling	Cloud experiments	Shows importance of metric selection for autoscaling.

NIST CSF 2.0 [5]	Cybersecurity governance	Framework and functions	Supports security governance alignment.
NIST AI RMF [7]	AI risk management	Govern-map-measure-manage orientation	Supports responsible AI control layer.

7. RESEARCH GAP

The reviewed literature has provided several gaps. To begin with, several AI-enabled cloud enterprise studies examine the model performance, with limited attention to enterprise governance integration. Besides, empirical validation is often done on different datasets, workloads, and baselines, creating limits on generalization. The interplay of cost, security and performance is the third issue, with mutual impact as one objective narrows, the other widens. Fourth, it still remains weak concerning the explanation of ML decisions on the cloud that trigger actions for automation. Few sector-specific frameworks exist for regulated industries. Many studies do not combine cloud workload traces with security data and business-process metrics. In this article, we address these gaps through a conceptual and analytical model merging resource optimization, predictive security, business automation and governance.

8. RESEARCH STRATEGY

The design is both analytical and secondary-data based. The research synthesizes a literature review, public dataset mapping, conceptual modelling, and analysis frameworks. The literature component consists of peer-reviewed and institutional sources related to cloud computing, autoscaling, workload prediction, AIOps, anomaly detection, cloud governance and AI risk management. The secondary-data included component detects public datasets that can be used for empirical validation. The paper does not state that they trained or evaluated ML models on the raw traces; they will put in the empirical values after validating the dataset.

The sources directly related to enterprise cloud computing, AI/ML-based resource management, cloud security analytics, AIOps, public cloud datasets, or cloud-governance frameworks were included. They were available through a reputable publisher or official institutional source, and they were relevant to the paper's research objectives. The criteria used for this review are exclusion criteria which include blog assertion which cannot be verified and promotional matter which does not contain any technical content. Another one was duplicate study, paper with no clear methodology and source which reported only numerical results without disclosable data or traceable experimental design.

The analytical dimensions cover issues such as resource utilization, workload predictability, autoscaling suitability, detection by security, alignment to governance, explainability, relevancen to cost control, SLA impact and business-automation relevance. The framework associates cloud telemetry with an AI/ML model, output of AI/ML with orchestration/security/business action and action to enterprise outcomes. The ethics include the protection of privacy, responsible management of AI, not producing fictitious results, transparency in automated decisions and human oversight for high-impact remediation action.

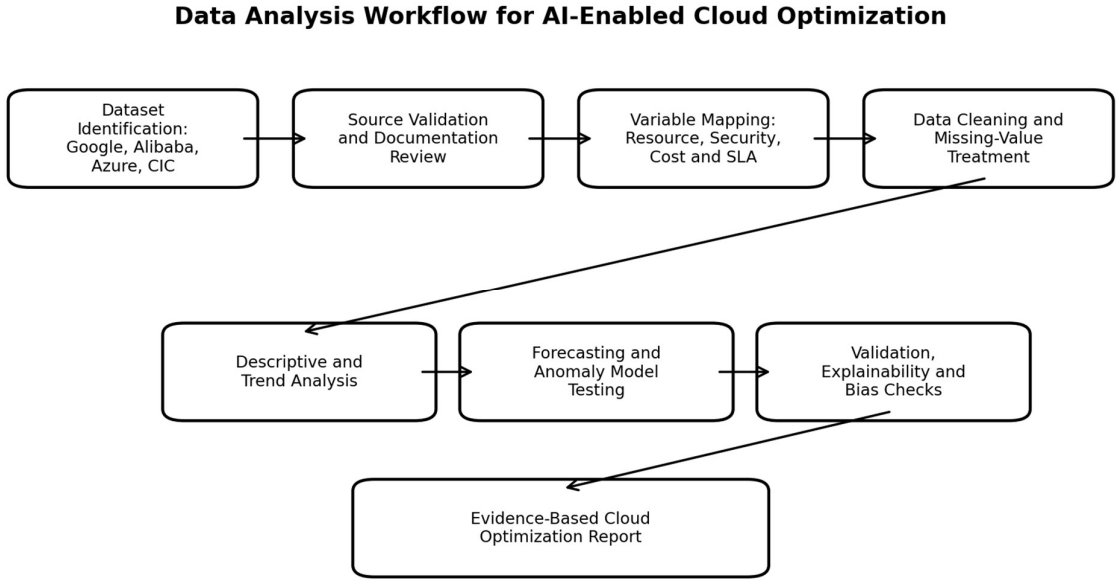

Fig. 2. Data analysis workflow for AI-enabled cloud optimization.

TABLE 3. AI AND ML TECHNIQUES USED IN ENTERPRISE CLOUD COMPUTING

Technique	Cloud Use	Strength	Caution
Regression	Capacity and demand estimation	Interpretable baseline	Weak for nonlinear behavior
Decision tree	Scaling rules and classification	Transparent rules	Overfitting risk
Random forest	Workload and security classification	Robust feature ranking	Less interpretable than a single tree
Neural network	Multivariate workload prediction	Captures nonlinear patterns	Requires large validated data
Reinforcement learning	Scheduling and autoscaling policy	Learns sequential decisions	Safe exploration is difficult
Clustering	Workload grouping and behavior profiling	Unsupervised discovery	Cluster meaning requires validation
Isolation Forest	Anomaly detection	Useful for sparse anomalies	Context-free alarms can be noisy
Autoencoder	Normal-behavior learning	Useful for complex telemetry	Explainability and drift issues

9. DATASET AND VARIABLES

The researchers selected public datasets to serve as potential reference empirical inputs for a number of reasons, most notably their traceability and reuse for further research. The load traces come from the Google cluster data generated from Borg-managed compute cells [9]. The Alibaba cluster dataset consists of production cluster traces such as the 2018 trace with around 4000 machines over eight days and the GPU trace with over 6500 GPUs on about 1800 machines over two months[10]. Microsoft Azure public traces comprise VM traces from the years 2017 and 2019, traces from Azure Functions, and inference traces from both Azure LLM and Azure multimodal. The CSE-CIC-IDS2018 dataset on AWS provides network traffic and system logs for labelled attack scenarios, and the extracted CICFlowMeter features [12].

Independent variables proposed are CPU Utilization, Memory Utilization, Storage Utilization, Network Traffic, Request Arrival Rate, Function Invocation rate, VM/container utilization, Scheduling class, Job duration, Login event, Failed login attempt, Anomalous traffic event and Configuration change. Dependent factors consist of resource wastage, SLA violation, scaling decision, predicted workload, anomaly label, incident severity, cost-risk category and automation response state. Factors like the type of workload, the time frame, the provider's environment, the geographical region, the criticality of the application, and the service-level goal can be the control variable. For cost modelling, it is important that verified billing data is provided by the enterprise or public benchmark; otherwise, data to be added after empirical verification.

TABLE 4. PUBLIC CLOUD DATASETS AND POSSIBLE ANALYTICAL VARIABLES

Dataset	Verified Source Information	Possible Variables	Use in This Study
Google Borg cluster traces	Borg-managed Google compute-cell workload traces [9]; 2019 trace analyzed across eight clusters for May 2019 [13].	Job events, task events, resource requests, scheduling class, failures	Workload prediction, scheduling and failure-risk framework
Alibaba cluster data	2018 trace: about 4000 machines for eight days; GPU trace: over 6500 GPUs on about 1800 machines for two months [10].	Machine usage, containers, batch tasks, DAGs, GPU workload variables	Co-located workload and MLaaS resource analysis
Azure Public Dataset	VM traces from 2017 and 2019; Azure Functions traces; LLM and multimodal inference traces [11].	VM lifetime, function invocation, blob access, token/image request variables	Serverless, VM packing and AI-inference workload analysis
CSE-CIC-IDS2018 on AWS	Seven attack scenarios; 50 attack machines; victim organization with 420 machines and 30 servers; 80 CICFlowMeter features [12].	Flow duration, ports, packets, bytes, labels, attack type	Predictive security and anomaly detection framework

TABLE 5. VARIABLES USED FOR CLOUD RESOURCE OPTIMIZATION AND PREDICTIVE SECURITY

Category	Variables	Possible Analysis	Data Status
Resource utilization	CPU, memory, storage, network I/O	Descriptive statistics, utilization distribution, underuse detection	Available in workload traces; empirical values to be inserted after validation
Workload demand	Request rate, job duration, invocation frequency	Trend analysis, time-series forecasting	Available in trace-specific schema
Autoscaling	Scaling decision, provisioning delay,	Policy comparison and SLA analysis	Requires implementation logs or simulation

	capacity change		
Security	Flow features, failed logins, attack labels, anomalies	Classification, anomaly detection, feature importance	Available in CSE-CIC-IDS2018; model values not computed here
Cost	Billing item, usage hours, discount plan, idle spend	FinOps analysis and rightsizing	Enterprise billing data to be inserted after empirical verification
Business automation	Ticket time, workflow status, decision latency	Process efficiency analysis	Enterprise process data to be inserted after empirical verification

10. DATA ANALYSIS AND RESULTS

The five layers are the conceptual framework. The initial component, known as the telemetry layer, gathers data on operational metrics, logs, traces, security events, identity events, billing records and SLIs. The data engineering layer is the second layer which cleans, labels, aggregates and normalizes heterogeneous data in the cloud. The AI/ML analytics layer essentially constructs predictive algorithms trained with historical data. It aids in job prediction, anomaly detection, capacity investigation, root-cause inference, cost-risk scoring, and even business-event classification. The intelligent orchestration layer performs autoscaling, scheduling, placement, incident response, policy enforcement and workflow automation are done in the fourth layer. The governance and decision-support layer in the fifth layer provides dashboards, audit trails, explainability, approval workflows, compliance evidence, and management reports.

The framework suggests that AI/ML models should not independently dictate or control high-risk enterprise actions. While modelling might suggest the necessity of scaling, blocking an identity, isolating a workload or changing a budget policy, the orchestration layer will evaluate more complex tasks. Specifically, it will consider service criticality, compliance obligations, rollback feasibility and human-approval thresholds. The NIST CSF has a substantial governance orientation, the NIST Zero Trust Architecture has substantial orientation and emphasis on resources and continuous verification, and the NIST AI RMF has a substantial focus on AI risk management.

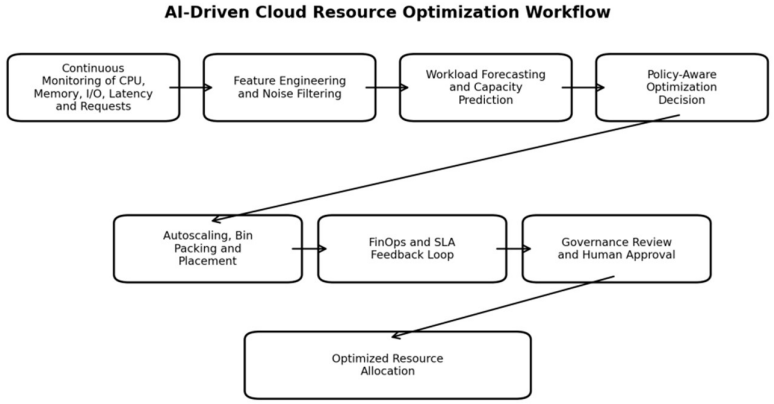

Fig. 3. AI-driven cloud resource optimization workflow.

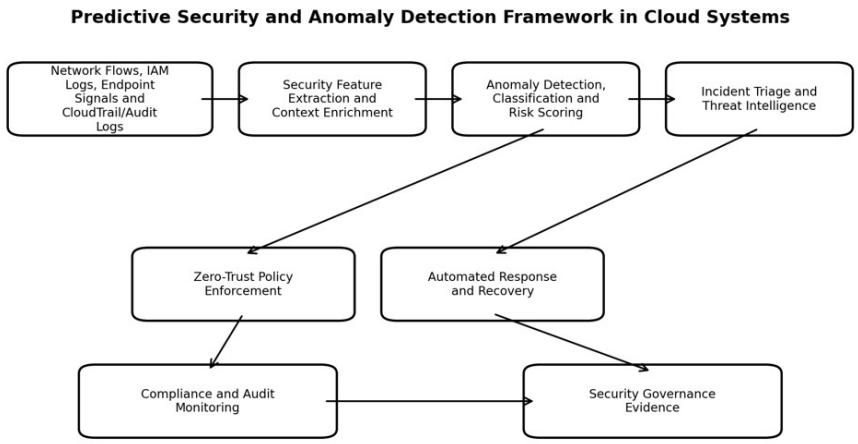

Fig. 4. Predictive security and anomaly detection framework in cloud systems.

11. CONCEPTUAL FRAMEWORK

The conceptual framework consists of five layers. The first

layer is the telemetry layer, which collects operational metrics, logs, traces, security events, identity events, billing records and service-level indicators. The second layer is the data engineering layer, which cleans, labels, aggregates and normalizes heterogeneous cloud data. The third layer is the AI/ML analytics layer, which performs workload prediction, anomaly detection, capacity forecasting, root-cause inference, cost-risk scoring and business-event classification. The fourth layer is the intelligent orchestration layer, which executes autoscaling, scheduling, placement, incident response, policy enforcement and workflow automation. The fifth layer is the governance and decision-support layer, which provides dashboards, audit trails, explainability, approval workflows, compliance evidence and management reports.

The framework assumes that AI/ML models should not directly control high-risk enterprise actions without governance checks. A model may recommend scaling, blocking an identity, isolating a workload or changing a budget policy, but the orchestration layer must evaluate service criticality, compliance obligations, rollback feasibility and human-approval thresholds. This design aligns with NIST CSF's governance orientation [5], zero-trust emphasis on resources and continuous verification [6], and NIST AI RMF's focus on AI risk management [7].

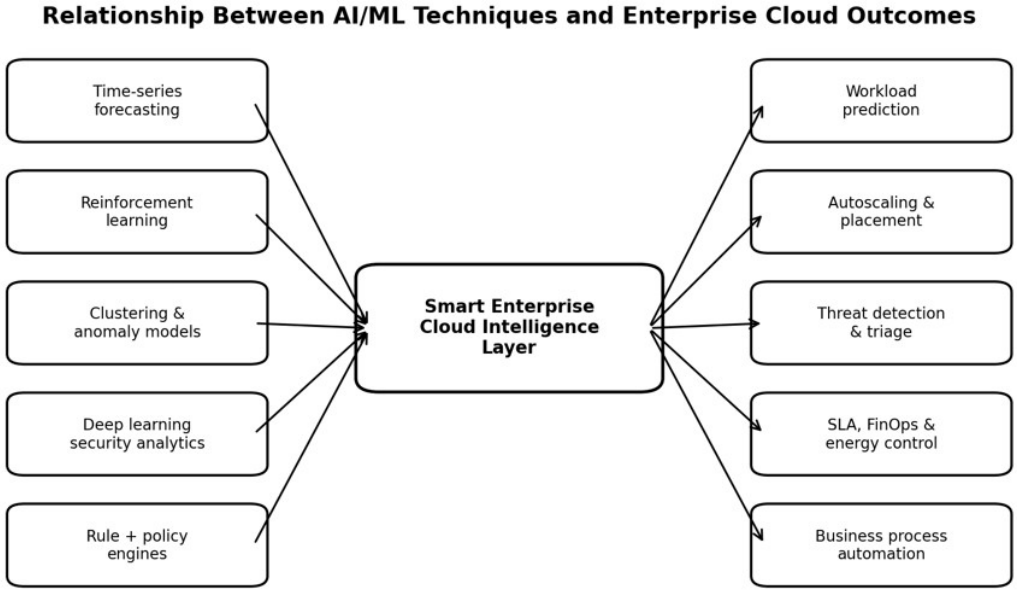

Fig. 5. Relationship between AI/ML techniques and enterprise cloud outcomes.

12. ANALYTICAL DISCUSSION

Regression models effectively predict resource demand using past information: request rate, CPU utilization, memory utilization and latency. While interpretable, the accuracy degrades when workloads behave non-linearly. Decision trees and random forests can model nonlinear relationships and provide feature-importance estimates, making them ideal for recommendations for autoscaling and security classification. Support vector machines can be effective for small structured datasets, but they tend to be challenging to scale for very large cloud telemetry streams.

In situations when workload behavior is multivariate and sequential or non-linear, deep learning or neural network models can be useful. LSTM and GRU variants can be applied to workload forecasting with recurrent neural networks, whereas autoencoders could learn normal representation for anomaly detection. Deep models need to be validated properly as high apparent accuracy can result from data leakage, class imbalance and poor generalization. Within production cloud operations, explainability, retraining cost, and drift monitoring are equally important as other predictive performance.

Reinforcement learning can learn a policy via interact with a simulated or controlled environment which makes it relevant for scheduling, autoscaling and placement. Safe exploration, bounded action, reward designing, and guardrails are basic requirements to use in enterprise clouds. An RL agent optimizing for cost may violate SLA targets due to the incompleteness of the reward function. On the contrary, a policy optimizing only the SLA is likely to overspend. As a result, the design of multi-objective rewards should consist cost and performance constraints for reliability, security as well as sustainability.

Models for clustering and anomaly detection aid in workload classification and security monitoring. Workload classes and abnormal patterns can be identified using k-means and DBSCAN clustering techniques. Deviations in resource uses or network behaviour can be detected using Isolation forest and autoencoder-based models. However, cloud aberrations are contextual. A product launch may be real because of spikes; privileged-access logs may be high risk because of smaller divergences. As a result, predictive security should merge ML scores with identity context, threat intelligence and zero-trust policies.

SLA Management; Artificial Intelligence/Machine Learning can predict breach risk before visibility of issues to customer. In terms of energy efficiency, workload consolidation and right-sizing may be supported by ML. Environmental claims should however be backed up by validated power and carbon data. Machine learning can help prioritize suspicious behavior in fraud and other threat detection, but security decision making needs response workflows with audit trails. In business automation, AI can generate recommendations for actions based on operational signals, for example, alerts about procurement, capacity planning, workflow routing or executive dashboards.

13. USES OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN CLOUD ENTERPRISE COMPUTING

One of the most straightforward applications is intelligent workload forecasting. Time-series models are able to predict demand for VMs, containers, databases, and serverless workloads. These predictions can be utilized by resource provisioning and autoscaling to provision capacity ahead of a spike in demand. Predictive maintenance identifies components prone to failure using traces, logs, and incidents. Maximizing Cloud Costs Using Usage Rightsizing Plan For Reserved Capacity Cleanup Incorrectly Tagged And Idle Resources Improves Chargeback.

Cybersecurity driven by artificial intelligence examines logs related to network traffic, identity, endpoint and audit cloud events. The worth lies not just in attack detection but also in alert overload reduction and risk-based incident prioritization. AI can be used for classification of sensitive data, detecting policy violations, monitoring of misconfigurations and generating audit evidence. Business intelligence and decision automation provide enterprise metrics from cloud telemetry to help managers assess the impact of cloud performance on customer experience, service revenue and process cycle time

Edge-cloud integration is becoming increasingly important for IoT, manufacturing, telecom and low-latency services. Based on latency, bandwidth, privacy, and cost constraints, machine learning (ML) can determine whether inference should take place at the edge or in the cloud. There are multiple aspects of the businesses where DevOps and AIOps automation can enhance deployment monitoring, incident response, and root-cause analysis. The maturity of these applications depends on standard telemetry, strong pipes, governance controls and workforce capability.

14. OBSTACLES AND DANGERS

Since cloud telemetry may have the identity attributes, customer metadata, IP addresses, application traces or regulated data, it raises a data privacy issue. Vendor lock-in can happen when AI automation is reliant on any of their tools, APIs or monitoring schemas. When we build models that learn from incomplete telemetry or imprecise historical operational decisions, we create problem of model bias and explainability. To maintain regulatory compliance, organizations should request their AI suppliers to offer evidence of auditability, access control, retention policy and that automated decisions are controlled.

The complexity of integration is high. Businesses frequently use cloud applications, on-premises systems, internet-based services and local headquarters for updates and management tasks. Integrating telemetry across these environments is a tough challenge both technologically and organizationally. The use of smart clouds may be impeded by a lack of required skills as cloud architecture is a requirement. Also engineering of data, engineering of ML, and security and governance against cyber attacks are required. If hackers can break into and attack the AI pipelines themselves through data poisoning, adversarial inputs, stolen user credentials or automation scripts, cybersecurity risks could greatly magnify.

Implementation Cost Must Be Evaluated Carefully Cloud management powered by AI needs an arsenal of data pipelines, monitoring tools, model hosting and retraining workflows. Plus, you need security controls and staff capability. Automated decisions that affect access, service continuation, customer prioritization or resource allocation raise ethical concerns. As a result of this, suitable smart enterprise cloud adoption should take up risk-tiered automation. Thus, low-risk actions can be automated, medium risk can be human-reviewed, while high-risk should be on-demand approval with rollback planning.

TABLE 6. CHALLENGES AND RISK-MITIGATION STRATEGIES

Challenge	Risk	Mitigation Strategy
Data privacy	Exposure of sensitive telemetry or customer metadata	Minimize collection, tokenize identifiers and enforce retention policy
Vendor lock-in	Dependence on proprietary tools and schemas	Use open telemetry, portable IaC and multicloud abstractions
Model opacity	Unexplainable scaling or security decisions	Use interpretable baselines, explainability reports and human review
Model drift	Declining prediction quality as workloads change	Monitor drift, retrain periodically and validate baselines
Cyberattack on automation	Poisoned telemetry or compromised scripts	Secure pipelines, least privilege and signed deployments
Compliance gaps	Automated actions violate regulatory obligations	Map policies to NIST CSF, AI RMF and sector requirements

Skill gaps	Poor model deployment and governance	Build cross-functional CloudOps, SecOps, DataOps and FinOps capability

15. PROPOSED SMART ENTERPRISE CLOUD COMPUTING MODEL

The proposed model is a governance-aware AI/ML cloud intelligence architecture. Its input layer receives workload telemetry, security logs, identity records, configuration states, billing records, application events and business-process indicators. The cloud data collection layer normalizes data through schema mapping, time synchronization, metadata tagging and privacy filtering. The AI/ML analytics layer includes forecasting, anomaly detection, classification, optimization, root-cause analysis and explainability modules. The intelligent orchestration layer translates model outputs into controlled actions such as scaling, right-sizing, scheduling, workload migration, access challenge, incident routing and workflow automation.

The predictive security layer applies zero-trust principles by treating identity, device, workload and data context as continuously evaluated signals. It supports threat detection, risk scoring, incident triage and automated containment while preserving audit logs. The business automation layer connects cloud events to enterprise workflows such as service management, procurement, compliance reporting, executive dashboards and customer-service triggers. The governance and compliance layer defines policies, approval thresholds, AI model review, privacy controls, regulatory mappings and human accountability.

Implementation should occur in stages. Stage 1 establishes observability, tagging, cost allocation and telemetry governance. Stage 2 develops descriptive dashboards and baseline rules. Stage 3 introduces ML forecasting and anomaly detection in advisory mode. Stage 4 deploys limited automation with rollback and human approval. Stage 5 expands automation to multicloud and business workflows with continuous monitoring. Stage 6 institutionalizes AI governance, model-drift detection and periodic audit. Expected benefits include improved resource awareness, more proactive security, better cost visibility, faster incident response and stronger decision support. These benefits require empirical validation and should not be treated as guaranteed numerical outcomes.

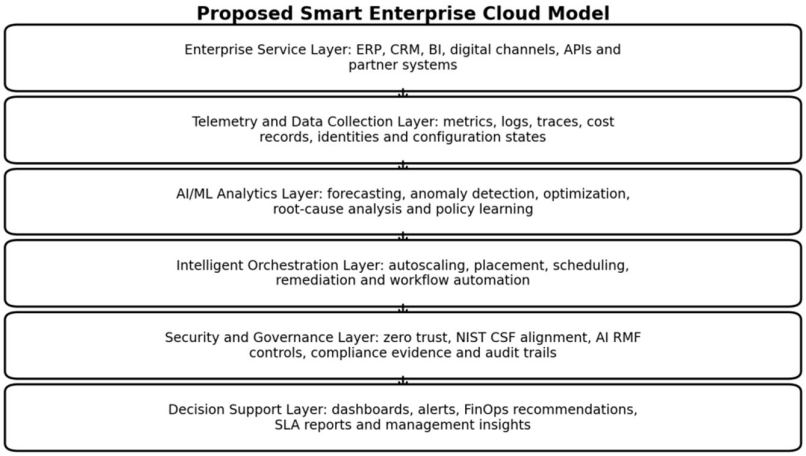

Fig. 6. Proposed smart enterprise cloud model.

TABLE 7. PROPOSED MODEL COMPONENTS AND FUNCTIONAL ROLES

Component	Function	Output	Governance Control
Telemetry layer	Collect metrics, logs, traces, cost and identity events	Unified cloud evidence base	Privacy filtering and retention policy
Data engineering	Clean, label, normalize and aggregate data	Model-ready dataset	Schema validation and lineage
AI/ML analytics	Forecast, classify, detect anomalies and optimize	Risk and action recommendations	Explainability, drift monitoring and testing
Orchestration	Scale, schedule, migrate and remediate	Controlled cloud action	Approval gates and rollback
Predictive security	Detect threats and enforce policy	Risk score and response trigger	Zero-trust policy and audit trail
Business automation	Route workflows and decision support	Operational and managerial action	Human oversight for high-impact decisions

TABLE 8. EVALUATION INDICATORS FOR SMART ENTERPRISE CLOUD IMPLEMENTATION

Outcome Area	Indicator	Measurement Approach	Notes
Resource optimization	Utilization, idle capacity, scaling frequency	Trace analysis and cloud-monitoring dashboards	Values require empirical dataset processing
Cost control	Idle spend, rightsizing recommendations, budget variance	Billing data and FinOps reports	Enterprise billing data required
Security	Detection latency, false positives, response time	Security logs and labelled datasets	Model results to be inserted after validation
Reliability	SLA violations, recovery time, error rate	SRE and incident-management records	Requires production/benchmark records
Business automation	Ticket cycle time, workflow completion, manual handoff	ITSM/process-mining data	Business data required
Governance	Auditability, explainability, policy compliance	Review logs and control testing	Qualitative and quantitative evidence

16. FINDINGS

According to studies, embedding artificial intelligence (AI) and machine learning (ML) in telemetry-rich, governance-aware operating models can support enterprise cloud computing. The study notes that workload forecasting, autoscaling and anomaly detection are among the most developed technical applications, while business automation, AI governance require further integration into the organization. Public datasets offer useful empirical pathways, but lack of comparability limits use for public benefit. The various cloud realities that Google, Alibaba, Azure and CSE-CIC sources refer to are production cluster workloads, co-located workloads, VM/serverless/AI inference traces and security attack scenarios [9]–[12].

Proposed framework can bridge forecasting models and orchestration policies to support intelligent resource allocation It may assist in predicting security through anomaly detection and zero-trust policy enforcement. It enhances operational efficiency by converting cloud telemetry into workflow and management actions. Before being able to make any claims on the reduction of cost, detection accuracy, SLA improvement or automation efficiency, empirical validation is needed. Consequently, the findings are conceptual and analytical, not experimental.

17. DISCUSSION

Cloud computing for smart enterprises can be seen as a capability for digital transformation, rather than a single technology. Value is created in AI and ML when they lead to improved operational decisions, security posture, cost discipline and business responsiveness.

To achieve this, cloud engineering, cyber security, data science, finance and business units must be aligned. Forecasts may be accurate but if the model does not change provisioning decision the enterprise value is limited. In the same vein, a security model without triage, escalation and remediation workflows that generates alerts may increase operational burden and not reduce risk.

The governance of cloud intelligence is also discussed herein. The NIST Cybersecurity Framework 2.0 addresses the necessity of governance and continuous security functions that demand accountability for risk decisions of AI-enabled cloud operations [5]. The concept of zero trust is important because perimeter assumptions cannot protect cloud assets and users. AI RMF is applicable to issues and risks raised by the cloud-control model as may affect organizational risk, privacy, reliability, and business continuity [7]. Collectively, these frameworks underpin a smart-cloud model which is automated, auditable, explainable and attuned to risk.

Operational performance should thus be assessed on multiple outcomes. The focus should be on calibrating rather than optimizing cost, performance, security, reliability, compliance and sustainability. Future work will need to test whether this framework can achieve this balance in validated datasets, reproducible pipelines and transparent baselines.

18. SUGGESTIONS

Enterprises need to focus on observability, tagging, identity management, and data-quality management. Telemetry exports, explainable recommendations and evidence of compliance should become more interoperable, say cloud service providers. IT managers must first use AI/ML models in advisory mode before granting permission to autonomously act. Cybersecurity teams have to combine ML anomaly detection with zero-trust controls, threat intelligence and incident response playbooks. Finance and procurement teams should align FinOps practices with resource-optimization models to ensure cost decisions reflect business value rather than only spend reduction.

Policymakers and regulators should put in place a framework for auditability and privacy-preserving telemetry in the cloud. Researchers must apply public traces responsibly, account for preprocessing decisions and avoid unverifiable claims. Cloud architects must design modular layers and composable smart-cloud system so the components related to forecasting, anomaly detection, orchestration, and governance are changing independently. No stakeholder should presume to replace governance with AI. Rather use my AI as a decision-supporting and automation-enhancing layer.

19. CONCLUSION WITH FUTURE SCOPE

This study presents a smart enterprise cloud computing

architecture and its applications using AI and ML techniques for intelligent resource optimization, predictive security and business automation. According to the analysis, AI/ML use cases are capable of improving enterprise cloud operations through forecasting, autoscaling, anomaly detection, root-cause analysis, FinOps support, compliance monitoring and workflow automation. The quality of these techniques is determined by telemetry quality, availability of validated datasets, appropriate governance controls, explainability and enterprise integration.

The smart enterprise cloud model that we propose brings together telemetry collection, data engineering, AI ML analytics, intelligent orchestration, predictive security, business automation and governance. A useful conceptual framework for enterprises seeking to evolve from merely reactive cloud management to adaptive, evidencebased and policyaware cloud operations. The manuscript does not make up numbers and points to public datasets that can enable future verification. The research can be done in the future to test the framework based on Google, Alibaba, Azure and security datasets empirically and building sector specific smart-cloud models. This includes developing explainable AI for cloud-related decision making and FinOps integration with AI driven optimization. Moreover, it evaluates edge-cloud intelligence, security AIOps frameworks design and implementation for regulated enterprise environment.

REFERENCES

1. P. Mell and T. Grance, "The NIST Definition of Cloud Computing," National Institute of Standards and Technology, Special Publication 800-145, 2011, doi: 10.6028/NIST.SP.800-145. Available: https://csrc.nist.gov/pubs/sp/800/145/final
2. Amazon Web Services, "AWS Well-Architected," AWS Architecture Center. Available: https://aws.amazon.com/architecture/well-architected/
3. Microsoft, "Azure Well-Architected Framework," Microsoft Learn. Available: https://learn.microsoft.com/en-us/azure/well-architected/
4. Google Cloud, "Google Cloud Well-Architected Framework," Google Cloud Architecture Center. Available: https://docs.cloud.google.com/architecture/framework
5. National Institute of Standards and Technology, "The NIST Cybersecurity Framework (CSF) 2.0," NIST CSWP 29, 2024, doi: 10.6028/NIST.CSWP.29. Available: https://www.nist.gov/cyberframework
6. S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," National Institute of Standards and Technology, Special Publication 800-207, 2020, doi: 10.6028/NIST.SP.800-207. Available: https://csrc.nist.gov/pubs/sp/800/207/final
7. National Institute of Standards and Technology, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, 2023, doi: 10.6028/NIST.AI.100-1. Available: https://www.nist.gov/itl/ai-risk-management-framework
8. Cloud Security Alliance, "Security Guidance for Critical Areas of Focus in Cloud Computing v5," 2024, updated 2025. Available: https://cloudsecurityalliance.org/artifacts/security-guidance-v5
9. [Google, "Borg cluster workload traces," GitHub repository. Available: https://github.com/google/cluster-data
10. Alibaba, "Cluster data collected from production clusters in Alibaba for cluster management research," GitHub repository. Available: https://github.com/alibaba/clusterdata
11. Microsoft Azure, "Azure Public Dataset: Microsoft Azure Traces," GitHub repository. Available: https://github.com/Azure/AzurePublicDataset
12. Canadian Institute for Cybersecurity, University of New Brunswick, "CSE-CIC-IDS2018 on AWS," dataset description. Available: https://www.unb.ca/cic/datasets/ids-2018.html
13. M. Tirmazi et al., "Borg: The Next Generation," in Proc. 15th European Conference on Computer Systems (EuroSys '20), ACM, 2020. Available: https://research.google/pubs/borg-the-next-generation/
14. M. Shahrad et al., "Serverless in the Wild: Characterizing and Optimizing the Serverless Workload at a Large Cloud Provider," in Proc. USENIX Annual Technical Conference, 2020. Available: https://www.usenix.org/conference/atc20/presentation/shahrad
15. Y. Cheng, Z. Chai, and A. Anwar, "Characterizing Co-located Datacenter Workloads: An Alibaba Case Study," arXiv:1808.02919, 2018. Available: https://arxiv.org/abs/1808.02919
16. Q. Cheng et al., "AI for IT Operations (AIOps) on Cloud Platforms: Reviews, Opportunities and Challenges," arXiv:2304.04661, 2023. Available: https://arxiv.org/abs/2304.04661
17. D. Saxena and A. K. Singh, "Workload Forecasting and Resource Management Models Based on Machine Learning for Cloud Computing Environments," arXiv:2106.15112, 2021. Available: https://arxiv.org/abs/2106.15112

18. I. Pintye, J. Kovács, and R. Lovas, "Enhancing Machine Learning-Based Autoscaling for Cloud Resource Orchestration," Journal of Grid Computing, 2024, doi: 10.1007/s10723-024-09783-1. Available: https://link.springer.com/article/10.1007/s10723-024-09783-1
19. M. Xu, C. Song, H. Wu, S. S. Gill, K. Ye, and C. Xu, "esDNN: Deep Neural Network Based Multivariate Workload Prediction in Cloud Computing Environments," ACM Transactions on Internet Technology, vol. 22, no. 3, 2022, doi: 10.1145/3524114.
20. M. Imdoukh, I. Ahmad, and M. Alfailakawi, "Machine Learning-Based Auto-Scaling for Containerized Applications," Neural Computing and Applications, vol. 32, pp. 9745-9760, 2020, doi: 10.1007/s00521-019-04507-z.
21. M. Straesser, J. von Kistowski, J. Grohmann, A. Eismann, A. Bauer, and S. Kounev, "Why Is It Not Solved Yet? Challenges for Production-Ready Autoscaling," in Proc. ACM/SPEC International Conference on Performance Engineering, 2022, doi: 10.1145/3489525.3511680.
22. L. Göcs and Z. C. Johanyák, "Identifying Relevant Features of CSE-CIC-IDS2018 Dataset for the Development of an Intrusion Detection System," arXiv:2307.11544, 2023. Available: https://arxiv.org/abs/2307.11544
23. Z. He and R. B. Lee, "CloudShield: Real-time Anomaly Detection in the Cloud," arXiv:2108.08977, 2021. Available: https://arxiv.org/abs/2108.08977
24. FinOps Foundation, "FinOps Framework," Linux Foundation project. Available: https://www.finops.org/
25. Singh, A., Bhushan, B., (2026) "AI and Machine Learning Enabled Enterprise Cloud Computing for Digital Transformation"" at Dandao Xuebao/Journal of Ballistics, **DOI:** https://doi.org/10.52783/dxjb.v38.333 , ISSN: 1004-499X. Vol. 38 No. 2, pp169 – 202