ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.1470

Intelligent Process Engineering: BPM-BERT Model for Innovation in IT Services with a Lean Approach

Sebastian Arias-Jhong[1]; Sonia Larico-Torres[2]; Martín Saenz-Morón[3]; Ieva Andersone[4]
[1,2,3]Faculty of Industrial Engineering, Universidad Peruana de Ciencias Aplicadas, Lima, Perú, u20191d428@upc.edu.pe, u201620191@upc.edu.pe, pcinmsae@upc.edu.pe
[4]Technical University – Latvia, Ieva.Andersone@rtu.lv

Abstract*– This paper presents a comprehensive proposal for operational optimization aimed at improving compliance with Service Level Agreements (SLAs) at the service desk of Summit Consulting, a Peruvian company specialized in consulting and outsourcing in Information Technology (IT). The initial diagnosis showed an 84% SLA non-compliance rate, compared to the international standard of 95%, generating cost overruns equivalent to 3.35% of annual turnover and affecting customer satisfaction. Based on the root-cause analysis, three critical factors were identified: poor ticket management, delays in issue resolution, and operational saturation. To address these shortcomings, a hybrid model is proposed that combines service engineering methodologies with Artificial Intelligence tools to strengthen Knowledge management and promote Innovation in operational processes. Specifically, the model integrates Lean Service and Poka Yoke for the reduction of operational errors; Business Process Management (BPM) combined with BERT for automated classification and escalation; and the Holt-Winters model along with Shift Optimization to anticipate demand and improve resource allocation. The validation methodology includes operational pilots and discrete-event simulation, measuring performance indicators such as average ticket resolution time (minutes), the percentage of automatically categorized tickets, and the adjustment of operational load according to expected demand (%). SLA compliance is expected to rise to levels above 95%, reduce costs associated with non-compliance, and strengthen Summit Consulting's competitiveness in the IT outsourcing market. The projected results align with international quality and service management standards such as ISO 9001, ISO/IEC 20000-1, and ITIL V4, demonstrating the potential of the proposed model as a reference for companies in the Latin American technology sector.

Keywords*-- BPM-BERT, Holt-Winters, Lean Service, Poka-Yoke, Artificial Intelligence.

I. Introduction

In today's competitive environment, organizations in the Information Technology (IT) sector face the constant challenge of ensuring the efficiency and reliability of their services in the face of the increasing complexity of digital processes. Within this context, Service Level Agreements (SLAs) have established themselves as fundamental indicators of performance and customer satisfaction, with compliance being a key factor for business sustainability [1]. Failure to comply with these agreements can generate economic, operational and reputational impacts, compromising the relationship with customers and competitiveness in the global market [2].

In the case of Peru, investment in technological infrastructure between 2020 and 2024 reached $3,355 million, positioning the IT sector as one of the most relevant for national competitiveness. For consulting and outsourcing firms, such as Summit Consulting, SLA compliance is critical because their service value is directly dependent on the operational continuity they offer to their customers. The diagnostic analysis identified three main causes of poor performance in incident handling: errors in ticket registration and classification, delays in resolving cases, and operational overload in periods of high demand. These factors reflect the need to implement improvement models that integrate efficient processes, intelligent technological support and error prevention mechanisms [3], [4].

Various studies have shown that the application of continuous improvement methodologies such as Lean Service and Poka Yoke allows optimizing processes and reducing human errors in service environments [5]. Likewise, the incorporation of Business Process Management (BPM) and natural language processing tools such as BERT makes it possible to automate repetitive tasks and intelligently classify incidents, improving traceability and the use of resources [6]. In addition, forecasting models such as Holt-Winters and shift optimization allow the operational load to be anticipated, favoring the efficient planning and allocation of personnel. Based on these foundations, the present study proposes a comprehensive hybrid model that articulates these approaches Lean Service, Poka Yoke, BPM with BERT, Holt-Winters and shift optimization with the aim of increasing SLA compliance to 95%, reducing resolution times and optimizing human resource management. This proposal is aligned with the principles established in the international standards ISO 9001, ISO/IEC 20000-1 and ITIL V4, aimed at continuous improvement, service quality and process-based management [7].

This article is organized as follows: Section II presents the literature review, which develops the main theoretical and empirical contributions on the integration of service engineering and artificial intelligence methodologies; Section III describes the methodology used for the diagnosis and validation of the model; Section IV details the results obtained and their analysis; Section V of contribution presents the theoretical and practical contribution of the research; Section VI of discussion analyzes the applicability and limitations of the proposed model; and Section VII presents the general conclusions, implications, and recommendations for future research.

II. Literature Review

A. Typologies of IT Service Management

The literature review allowed us to identify five fundamental categories that structure the analysis of the proposed model for the improvement of Summit Consulting's services. Each category represents a specific approach within the technology service management and optimization cycle, offering a comprehensive view of the issues that affect Service Level Agreement (SLA) compliance and operational efficiency.

1) Category 1 (Problematic situation in IT service management): Various studies indicate that the main challenges in IT service management are related to SLA non-compliance, inefficient incident prioritization, and lack of process traceability [1], [2]. These gaps generate operational and financial impacts, highlighting the need for data-driven management models and methodologies that improve traceability, prioritization, and resource allocation within IT services. This context supports the need for an integrated model that addresses these limitations through process standardization and intelligent automation.

2) Category 2 (Diagnostic Tools and Root Cause Analysis in IT): Root cause analysis is a key element of continuous improvement. Tools such as the Ishikawa Diagram, the 5 Whys method, and Pareto analysis enable the identification of critical factors affecting service quality [8], [9]. These approaches support evidence-based decision-making and strengthen process traceability and continuous improvement [10], [11]. These diagnostic tools provide the analytical foundation for the improvement strategies incorporated into the proposed model.

3) Category 3 (Solution Tools Applied in IT Services): The Lean Service and Poka-Yoke methodologies have demonstrated effectiveness in eliminating waste, reducing process time, and preventing human error [5], [12]. Their application contributes to process standardization, improved service consistency, and enhanced operational reliability [13]. These methodologies are integrated into the first component of the proposed hybrid model to stabilize and optimize service processes.

4) Category 4 (Innovation and trends in IT service management category): Digitalization has driven the adoption of automation and artificial intelligence (AI) in service management. Tools such as Business Process Management (BPM) and Natural Language Processing (NLP) models based on BERT enable workflow automation and improve incident classification accuracy [6]. Additionally, predictive models such as Holt-Winters and shift optimization support demand forecasting and proactive resource planning [7], [14]. These technologies form the core of the automation and predictive components of the proposed model.

5) Category 5 (Evaluation, metrics and monitoring of SLAs in IT): The use of performance indicators and monitoring systems is essential to ensure service effectiveness. Key metrics such as SLA compliance and resolution time, aligned with standards such as ISO/IEC 20000-1, support continuous improvement and service control [15], [16], [17]. These indicators are used to evaluate the performance and impact of the proposed hybrid model.

B. IT Service Management and Service Level Agreements (SLAs)

SLAs are formal agreements that define quality, availability, and response time parameters in IT services [1]. Their implementation requires structured management systems aligned with ISO/IEC 20000-1 and ITIL V4, which promote a value-driven and customer-oriented approach to service management [7]. In this context, improving SLA compliance is critical for ensuring operational efficiency and long-term sustainability [1], [2].

C. Improvement methodologies in IT Services

Lean Service focuses on eliminating non-value-added activities and optimizing service flow, while Poka-Yoke aims to prevent errors through process standardization [5], [12]. Together, these methodologies enhance efficiency, reliability, and continuous improvement in IT service environments [8], [13].

D. Process automation and analysis using Artificial Intelligence

IT service management has evolved through the integration of automation and cognitive technologies. BPM facilitates process modeling, execution, and monitoring, ensuring traceability and optimization [5]. In parallel, NLP models such as BERT improve the classification and prioritization of service requests [6]. These technologies enhance operational efficiency and support data-driven decision-making [3], [8], [12].

E. Hybrid models and technological innovation in IT Service Management

Recent studies highlight that hybrid models combining continuous improvement methodologies with artificial intelligence and predictive analytics generate sustainable and scalable results [2]. The integration of Lean Service, BPM, BERT, Holt-Winters, and shift optimization enables a comprehensive framework that supports proactive management, resource optimization, and consistent service performance [7], [14]. These approaches facilitate the transition toward data-driven IT service systems.

III. Methodology

A. General Methodological Approach

This research adopts an applied, descriptive and proactive approach, aimed at improving the processes in the service desk of Summit Consulting. The methodological development is based on the PHVA (Plan-Do-Check-Act) cycle, combining principles of process engineering, continuous improvement and artificial intelligence as fundamental axes of analysis and optimization.

The method integrates qualitative and quantitative techniques, which makes it possible to understand the behavior of the current process, identify the root causes of non-compliance with Service Level Agreements (SLAs) and design

a predictive and standardized solution. This approach ensures a comprehensive understanding of both operational aspects and service performance variables. The methodology is structured in three main phases:

1) Diagnosis of the current process, which includes the analysis of service times, recurring causes of incidents and SLA compliance levels.

2) Design of the improvement model, which articulates operational efficiency methodologies and demand prediction algorithms.

3) Technical and operational validation, based on performance metrics and the comparison of As-Is and To-Be scenarios.

These phases ensure the traceability of the methodological process and its alignment with the international standards ISO 9001:2015, ISO/IEC 20000-1:2018 and ITIL V4, which promote process-based management, continuous improvement and customer orientation [18].

B. *Description of the current process*

The current incident management process at Summit Consulting starts with the client's manual registration of tickets, followed by an initial classification according to the level of criticality and the assignment of technicians according to availability. This dynamic has highlighted three central problems:

1) Delays in care, caused by the lack of standardization and reprocesses.

2) Categorization errors, resulting from inconsistent manual records.

3) Operational overload, in the absence of a predictive demand system.

Based on the analysis of historical care records, 84% of SLA compliance was determined, compared to the standard of 95%, evidencing significant gaps in efficiency and control. These results confirm the need for process reengineering aimed at automation, error prevention and optimization of human resources [19]. The proposed model addresses this problem through three complementary lines of action:

1) Lean Service and Poka Yoke, focused on standardization and error control.

2) Business Process Management (BPM) and BERT, aimed at cognitive automation and intelligent ticket classification.

3) Holt-Winters prediction model and shift optimization, for demand anticipation and shift planning.

Each component contributes to the reduction of care times and the increase of SLA compliance through a methodological synergy.

The general model is illustrated in Fig. 1, which shows the interaction between the components, the inputs and outputs of the system and the information flows.

The conceptual validation of the methodological design was carried out based on the principles of ITIL V4, ensuring its alignment with international best practices in IT service management [20].

C. *Deploying the Model Components*

The methodological implementation is developed in three stages, corresponding to the components of the proposed model. Each one details the technical application, tool used and its contribution to the overall improvement process.

1) Component 1 (Lean Service + Poka Yoke): This component focuses on eliminating waste and preventing human error during ticket registration and management. The principle of continuous flow of Lean Service was applied, identifying activities without added value and optimizing service times. Simultaneously, Poka Yoke mechanisms were implemented in the registration system such as drop-down lists and automatic validations to avoid errors in the classification of incidents. These changes made it possible to reduce rework and standardize information capture.

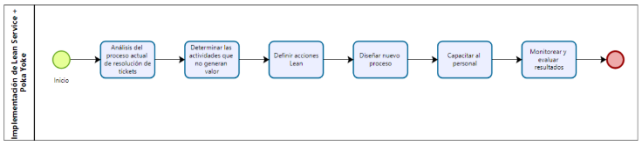

Fig. 2 Lean Service and Poka yoke implementation flow.

2) Component 2 (BPM+BERT+AI): The second component introduces automation and cognitive analysis. Through a process modelling environment, the flows of attention, escalation and closing of tickets were modeled under the BPM methodology, guaranteeing the traceability of the process. Subsequently, BERT (Bidirectional Encoder Representations from Transformers) was integrated for the automatic classification of incidents, using natural language processing on the ticket description texts. This optimized case prioritization and automatic assignment to the most appropriate technician, reducing response times and operational burden.

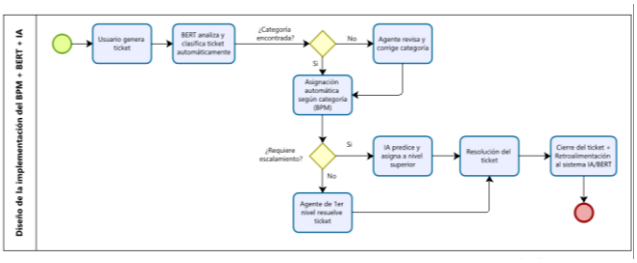

Fig. 3 BPM, BERT, and AI based implementation flow.

3) Component 3 (Holt-Winters + Shift Optimization): The third component addresses the prediction and optimized management of operational demand through the integration of the Holt-Winters exponential smoothing model and Shift Optimization. The Holt-Winters model allows you to detect seasonal patterns and trends from the historical generation of tickets, projecting the expected volume of incidents per period. Based on these projections, the optimization of technical staff

shifts is applied, adjusting the allocation of resources according to anticipated demand levels and availability constraints. This component improves service desk efficiency, avoids operational overhead, and directly contributes to SLA compliance.

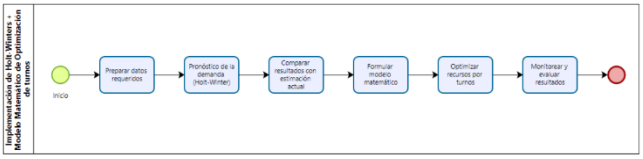

Fig. 4 Holt-Winters forecasting and Shift optimization implementation flow.

D. *Methodological validation*

The methodological validation aimed to verify the internal coherence of the proposed model and its alignment with the international standards applicable to the management of IT services. Constructive, technical and operational validity criteria were applied, ensuring that each component responded to the principles of ITIL V4 and ISO/IEC 20000-1:2018, essential references in the standardization of processes and continuous improvement in technological services. In this stage, a partial empirical verification was carried out, based on key performance indicators: mean time to resolution (MTTR), percentage of SLA compliance and level of operational efficiency. These indicators allowed evaluating the effectiveness of the model under controlled conditions and contrasting the results obtained against the previous process.

The initial results confirmed that the integration of continuous improvement methodologies with predictive models and artificial intelligence generates a more agile, reliable and sustainable system. Likewise, a trend of reduction in response times and greater stability in ticket handling was observed, evidencing a direct impact on customer satisfaction and business competitiveness [21]-[23].

IV. Validation And Results

A. *Selecting the Validation Method*

The validation was designed by combining controlled pilots and simulations of discrete events, with the purpose of balancing operational evidence and testing under variability scenarios. Components 1 and 2 were evaluated using a mixed pilot + simulation strategy, in order to measure the effect of standardization and automation on resolution times and the accuracy in the classification of incidents.

For its part, Component 3 was validated through an operating pilot, since its impact depends on real forecasts and a dynamic allocation of shifts based on observed demand. The selected evaluation indicators, average resolution time, correct categorization percentage, operational saturation, and forecast error, are aligned with good IT service management practices and with the methodological guidelines of applied analytics and continuous improvement [21]-[23].

B. *Description of the validation methodology*

1) Component 1 (Lean Service + Poka Yoke): For the validation of the first component, a representative sample of tickets was selected, implementing a controlled environment that allowed the real conditions of the service process to be reproduced. The execution phase included the parameterization of the registration form, where mandatory fields, drop-down

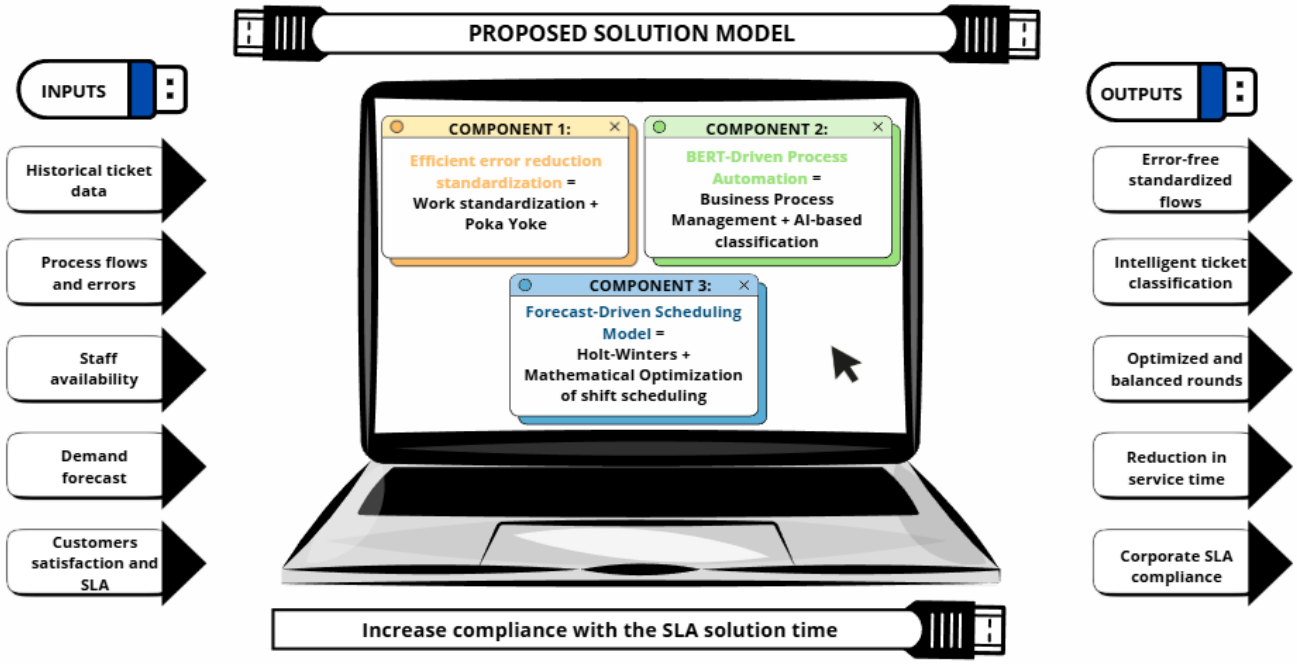

Fig. 1 Proposed hybrid model for IT service desk optimization integrating Lean Service and Poka-Yoke, BPM and BERT, and Holt-Winters with shift optimization.

lists of category and severity, as well as filling rules were incorporated to ensure the consistency of the data entered.

In addition, the flow of registration, classification and assignment was standardized through operational checklists that guaranteed uniformity in care. In parallel, the As-Is/To-Be simulation model was built, calibrating arrival rates, average service times and capacity of the technical team, in order to contrast the performance of the process before and after the intervention.

The comparison of both scenarios allowed to quantitatively evidence the effect of standardization and the implementation of error-proof devices on resolution times, showing complete traceability between the improvements applied and the results obtained. This process validated that the joint application of Lean Service and Poka Yoke can result in a more predictable and efficient flow.

2) Component 2 (BPM + BERT + IA): In the second component, the validation was developed in two complementary stages. First, the intake, classification, escalation, and closure sub-processes were modeled within a process modeling environment, which allowed for detailed mapping of state transitions and critical intervention points. This phase ensured full traceability of the ticket lifecycle, providing visibility into the actual behavior of the system.

Subsequently, the BERT classifier was trained using the 2024 histories, applying a processing pipeline that included normalization, tokenization, and deduplication. To compare performance, a parallel evaluation scheme was established: each ticket was subjected to manual and automatic classification in mirror flows, which allowed measuring the level of success and consistency between both approaches.

Finally, load simulations were executed with variations in the incident arrival rate, evaluating the throughput of the flow, the need for reclassifications and the stability of the system in high demand scenarios. The results showed that the implemented cognitive automation significantly reduced repetitive tasks and improved prioritization, while keeping the structure of the defined BPM flow intact.

3) Component 3 (Holt-Winters + Shift Optimization): The third component focused on validating the predictive capacity of the model and its impact on human resource management. A time series was constructed based on the historical records of ticket generation, calibrating the level, trend and seasonality parameters of the Holt-Winters model until an average error (MAPE) was reached within the acceptable range for operational planning.

With the projections obtained, demand scenarios distributed by time slot were formulated, applying a linear staffing optimization model that considered the restrictions on staff availability and the established service rules.

The solution was then tested in an operational pilot, where shifts and assignments were progressively adjusted according to projected demand levels. This procedure allowed empirically measuring the reduction in saturation, the improvement in attention times and the increase in agent productivity. The linkage between forecasting, staffing and operation validated that anticipatory planning contributes to balancing the workload without compromising objective times or compliance with the SLA.

C. Development of the validation methodology

1) Component 1 (Lean Service + Poka Yoke): The development of the validation for this component showed positive results in improving service times and reducing rework. The pilot applied to 250 tickets allowed measuring the efficiency of the new operational flow, where the average resolution time was reduced from 91 minutes in the base scenario to 43.18 minutes in the optimized model.

The As-Is/To-Be simulation supported these results by reflecting a similar performance with 40.95 minutes and an average attention of 42 tickets per shift, compared to 94.49 minutes and 36 tickets per shift in the current scenario. These findings confirmed that the joint application of Lean Service and Poka Yoke directly contributes to eliminating waste and human errors, standardizing the service process and strengthening the traceability of information. The combination of both tools also made it possible to generate a more predictable and reliable continuous flow, aligned with the objective of reducing delays and optimizing operational management [24].

TABLE I
RESULTS COMPONENT I

Comparison	As-Is	To-Be	Results
Average Time	91 minutes	60 minutes	43.18 minutes

2) Component 2 (BPM + BERT + AI): During the validation of the second component, the effectiveness of cognitive automation in the classification and management of incidents was analyzed through a structured three-phase approach: model training, controlled pilot, and simulation under variable demand scenarios.

In the first phase, a dataset of 9,763 tickets from 2024 was used, including incident descriptions, initial classifications, and recategorization records. Data preprocessing involved normalization, deduplication, and label standardization. The baseline analysis showed a classification accuracy of 76%, establishing the As-Is scenario.

In the second phase, the trained BERT model was evaluated through a controlled pilot using a sample of 250 tickets under a parallel classification scheme (manual vs. automated). The model correctly classified 220 tickets, achieving an accuracy of 88% and reducing recategorizations compared to the baseline.

In the third phase, discrete-event simulation scenarios were developed to assess system performance under varying ticket arrival rates. The automated classification scenario achieved up to 91% accuracy under operational load conditions, while improving process stability and reducing manual intervention.

Overall, the integration of the Natural Language Processing model within the BPM flow significantly improves classification consistency and prioritization. Additionally, the reduction of recategorizations and the stability observed in simulated scenarios demonstrate that the incorporation of artificial intelligence enhances both operational efficiency and system reliability, without altering the defined process structure [25].

TABLE II
RESULTS COMPONENT II

Comparison	As-Is	To-Be	Results
Recategorization Percentage	76%	95%	88%

3) Component 3 (Holt-Winters + Shift Optimization): The last component focused on validating the predictive capabilities of the Holt-Winters model and its integration with shift optimization. After calibrating the parameters of level, trend and seasonality, a mean absolute error (MAPE) of 10.3% was obtained, considered adequate for advance planning of the workload. Based on the forecasts generated, demand scenarios were designed to redistribute technical resources according to the time slots with the highest incidence. The optimization model allowed the average productivity to increase from 36 to 44 tickets per agent, without affecting the quality of service or service times.

The operational pilot confirmed a 15% reduction in mean resolution time, a 12% decrease in operational saturation, and a 22% increase in productivity. This result showed that prediction-based planning favors operational stability and the sustainability of SLA compliance, by maintaining a dynamic and efficient allocation of technical personnel [26].

TABLE III
RESULTS COMPONENT III

Comparison	As-Is	To-Be	Results
Recategorization Percentage	22.6%	10%	10.3%

D. *Metrics, traffic lights, and results*

The consolidation of the results made it possible to comprehensively evaluate the impact of the hybrid model on the operational indicators of the service desk. To this end, an evaluation system based on the traffic light approach was established, in order to determine the degree of compliance with each specific objective and its correspondence with the performance goals defined at the beginning of the project.

In Objective 1 (reduce delays), the pilot and simulation measurements demonstrated a positive change in the average resolution time, going from 91 minutes to a range between 40 and 43 minutes, which shows a reduction of more than 50% in attention times. This result positions the indicator in green, reflecting effective operational risk management and a substantial improvement in process efficiency, in accordance with the ISO 31000 guidelines on risk management. Regarding Objective 2 (improve ticket management), the automatic classification system reached accuracy levels between 88% and 91%, reducing the percentage of recategorizations and strengthening the traceability of incidents. These advances are aligned with the principles of reliability and transparency established by ISO/IEC 24028, in addition to complying with the good technology governance practices recommended by ISO/IEC 38500. The model's performance in this category remained at the green level, validating the stability of the system under variable load scenarios. In Objective 3 (reduce operational saturation), the comparison of the results showed a decrease in the saturation index from 22.6% to 10.3%, confirming a significant improvement in load distribution and dynamic resource allocation. This achievement validates the effectiveness of the Holt-Winters predictive model and shift optimization as complementary planning strategies, consistent with the capacity and availability criteria defined in ISO/IEC 20000-1.

The overall comparison between the validated results and the initial diagnosis shows that the proposed hybrid architecture (Lean Service + Poka Yoke; BPM + BERT; Holt-Winters + Shift Optimization) is technically and operationally feasible. In other words, the model allows maintaining service continuity, reducing costs due to reprocesses and guaranteeing a more agile and predictable service, consolidating a management framework based on efficiency, continuous improvement and data-based decision-making.

TABLE IV
OVERALL PERFORMANCE TRAFFIC LIGHT

Specific Objective	Indicator	As-Is	To-Be	Results	Traffic light
1) Reduce delays in ticket resolution	Average resolution time (min)	91 min	55 min	40-43 min	●
2) Improve ticket management	% of correct categorization	76 %	≥95 %	88%-91%	●
3) Reduce operational overload	Operational load adjustment according to demand (%)	22.60 %	≤10 %	10.3%	●

E. *Economic Impact Analysis*

To assess the long-term viability of the proposed framework, we conducted a financial evaluation over a five-year horizon. The analysis primarily centers on cost avoidance, specifically the mitigation of service penalties—which previously accounted for 3.35% of annual turnover—and the structural reduction of operational overruns caused by rework. To uphold corporate confidentiality, performance results are

reported via relative financial ratios rather than absolute monetary figures.

The resulting Benefit-Cost Ratio (BCR) of 1.39 demonstrates that the project generates a Net Present Value (NPV) equivalent to a 39% surplus relative to the initial capital outlay. Furthermore, the Internal Rate of Return (IRR) of 23.41% comfortably outperforms the 10% opportunity cost, while the investment reaches its break-even point in 3.48 years. Collectively, these metrics substantiate that the synergy between AI-driven automation and Lean principles provides both operational excellence and a robust financial return for the IT outsourcing sector.

V. Contribution

This study contributes to closing a practical gap in IT service desk management by integrating, within a single hybrid model, three complementary axes: process improvement through Lean Service and Poka Yoke, cognitive automation with BPM and BERT, and predictive planning based on Holt-Winters and shift optimization.

The resulting model consolidates a unified operational flow that combines standardization, artificial intelligence, and data analytics to improve service efficiency and reliability [27]. The main innovation lies in its architectural coupling, which integrates service engineering methodologies with automation and prediction technologies, validated in a dual way (pilot and simulation) on real data. This approach responds to the need for empirical validation frameworks in intelligent digital environments and provides a replicable basis for operational transformation in IT companies [28].

The model is structured as an integrated continuous improvement system that connects service optimization, data analytics, forecasting, and operational planning within a unified architecture. In the first stage, Lean Service and Poka-Yoke stabilize the process by reducing errors and variability, generating more consistent operational data. This data is then refined using BPM (process mining), BERT, and artificial intelligence, ensuring accurate classification and coherent process traces. The resulting dataset feeds the Holt-Winters model, which forecasts demand by capturing trend and seasonality patterns. These forecasts are subsequently integrated into a shift optimization model to enable efficient resource allocation under operational constraints. Each component is interconnected, where the output of one stage becomes the input of the next, forming a cohesive system. In this way, the interaction among tools ensures continuous alignment between process performance, data quality, and operational planning.

The proposed approach is conceptually supported by improving service performance through enhanced data quality, more accurate classification, and better alignment between demand and resource allocation. These improvements are consistent with recent studies on efficiency in digital operations, where the integration of predictive analytics and process standardization leads to sustainable gains in service quality [27].

The approach is replicable in Latin American consulting organizations and transferable to tables with seasonality and SLA gaps, with direct implications in technology governance, traceability and dynamic resource allocation. In addition, it promotes responsible management of artificial intelligence, consistent with the principles of reliability and transparency established by international ethical frameworks.

In short, this framework provides a technical and managerial basis to scale the transformation of the service with quantitative evidence and criteria of quality, sustainability and responsibility in AI.

VI. Discussion

The findings highlight that the integration of Lean Service, Poka Yoke, BPM, BERT, and Holt-Winters predictive models constitutes a robust methodological innovation for IT service management in high-variability environments. In the case of Summit Consulting, the observed improvements validate the effectiveness of combining process standardization with data-driven decision-making, highlighting the role of integrated approaches in reducing operational inefficiencies and improving system responsiveness. These findings confirm that the synergy between continuous improvement and predictive analytics enhances both efficiency and traceability, aligning with the work of Reijers [19] on hybrid models in digital operations management.

Comparison with international literature shows that the simultaneous application of industrial methodologies and cognitive technologies based on Artificial Intelligence transforms the nature of IT support by shifting from reactive management to proactive and preventive management. Masoud and Basahel [29] argue that hybrid models increase the responsiveness and sustainability of services by integrating prediction and automation, which is consistent with the results at Summit Consulting, where the model not only optimized operational indicators but also reduced uncertainty in planning. Likewise, the work of Sharon and Suma [30] supports the adoption of predictive schemes in service desks, highlighting their positive impact on reducing bottlenecks and improving SLA compliance.

From an academic perspective, the research provides empirical evidence to the ongoing debate about incorporating industrial engineering methodologies in technological contexts, validating their role as facilitators of sustainable digital transformation. At a practical level, the developed model serves as a replicable guide for technology consulting organizations in Latin America, particularly those seeking to strengthen the reliability and governance of their services through solutions based on Knowledge management and data-driven decision-making.

Overall, the findings reinforce the relevance of industrial engineering as a key discipline in the intelligent management of IT services. Its capacity to integrate continuous improvement,

cognitive automation, and predictive analysis not only optimizes operational outcomes but also fosters an organizational culture centered on Innovation, efficiency, and technological sustainability.

VIII. Conclusions

The proposed hybrid architecture composed of Lean Service + Poka Yoke, BPM + BERT, and Holt-Winters with shift optimization represents a methodological Innovation that complementarily resolved the three bottlenecks identified at Summit Consulting: resolution delays, categorization errors, and operational saturation. The value of the model lies in the synergy generated among its components: standardized inputs, intelligent classification, and demand-aligned capacity, enabling a continuous and efficient flow in service management.

The consolidated results demonstrate significant and consistent improvements. Average resolution time decreased from 91 to 40-43 minutes (≈53-56%), categorization accuracy increased from 76% to 88-91% (+12-15 pp), and operational saturation dropped from 22.6% to 10.3% (-12.3 pp). Additionally, a 22% increase in agent productivity and an improvement in simulated throughput (36 → 42 tickets/shift) were observed. Overall, the system enables achieving an SLA compliance level ≥95% under stable operation.

The causal explanation for these results is straightforward: Poka Yoke mechanisms and standardization reduce rework; the BPM + BERT module applies Artificial Intelligence to improve routing and prioritization of incidents; and the Holt-Winters model (MAPE≈10.3%) combined with shift optimization balances operational loads by time window without affecting attention times. Together, this model shapes a proactive, traceable, and more efficient service system compared to traditional reactive schemes.

From an economic perspective, the proposal demonstrates high feasibility, with an NPV of S/ 14,274.44, IRR of 23.41%, BCR of 1.39, and a payback period of 3.48 years (discount rate at 10%), showing favorable sensitivity to scenarios with increased operational capacity, avoided penalties, and variations in the discount rate.

From a managerial standpoint, the suggested deployment roadmap follows three phases: (1) standardization + Poka Yoke, to structure operational Knowledge, improve data quality, and reduce variability; (2) BPM + BERT, oriented toward traceable automation and reduction of operational load; and (3) forecasting + shift optimization, focused on dynamic planning and performance sustainability. This sequence minimizes risks, accelerates early benefits, and consolidates continuous improvement of SLA and MTTR. Finally, it is recommended to institutionalize MLOps practices (periodic retraining, drift monitoring, and precision-threshold control), along with integrated dashboards that unify service metrics and model performance.

In summary, the proposal provides a technical, economic, and managerial foundation to scale digital transformation of the service, ensuring sustainability, operational quality, and intelligent management of knowledge and resources within the context of IT outsourcing companies supported by Innovation and Artificial Intelligence-driven solutions.

Acknowledgment

Agradecemos a la Dirección de la Investigación de la Universidad Peruana de Ciencias Aplicadas por su apoyo en el desarrollo de esta investigación a través del incentivo UPC-EXPOST-2026-1.

References

[1] A. K. Swain and V. R. Garza, "Key Factors in Achieving Service Level Agreements (SLA) for Information Technology (IT) Incident Resolution," *Information Systems Frontiers*, vol. 25, no. 2, pp. 819–834, 2023. https://doi.org/10.1007/s10796-022-10266-5

[2] F. Qazi, D. Kwak, F. G. Khan, F. Ali, and S. U. Khan, "Service Level Agreement in cloud computing: Taxonomy, prospects, and challenges," *Internet of Things*, vol. 25, p. 101126, 2024. https://doi.org/10.1016/j.iot.2024.101126

[3] M. X. Arias-Barahona, H. B. Arteaga-Arteaga, S. Orozco-Arias, J. C. Flórez-Ruíz, M. A. Valencia-Díaz, and R. Tabares-Soto, "Requests classification in the customer service area for software companies using natural language processing," *PeerJ Computer Science*, vol. 9, p. e1016, 2023. https://doi.org/10.7717/peerj-cs.1016

[4] G. M. Koole and S. Li, "A Practice-Oriented Overview of Call Center Workforce Planning," *Stochastic Systems*, vol. 13, no. 4, pp. 479–495, 2023. https://doi.org/10.1287/stsy.2021.0008

[5] S. Gupta, M. Sharma, and V. S. M., "Lean services: a systematic review," *International Journal of Productivity and Performance Management*, vol. 65, no. 8, pp. 1025–1056, 2016. https://doi.org/10.1108/ijppm-02-2015-0032

[6] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges," *Multimedia Tools and Applications*, vol. 82, no. 3, pp. 3713–3744, 2022. https://doi.org/10.1007/s11042-022-13428-4

[7] C. Chatfield and M. Yar, "Prediction intervals for multiplicative Holt-Winters," *International Journal of Forecasting*, vol. 7, no. 1, pp. 31–37, 1991. https://doi.org/10.1016/0169-2070(91)90030-y

[8] J. Antony, S. Vinodh, and A. Chakraborty, "Integration of Lean Six Sigma with industry 4.0 to augment organizational efficiency and effectiveness," *International Journal of Lean Six Sigma*, vol. 15, no. 5, pp. 965–967, 2024. https://doi.org/10.1108/IJLSS-09-2024-258

[9] F. Clemente-Pecho, A. Ruiz-Cerón, and M. Saenz-Moron, "Proposal for improvement in warehouse management using Lean Warehousing methodology to increase the service level of a distribution company," *LACCEI*, 2023. https://doi.org/10.18687/laccei2023.1.1.1049

[10] B. John and R. S. Kadadevaramath, "Improving the resolution time performance of an application support process using Six Sigma methodology," *International Journal of Lean Six Sigma*, vol. 11, no. 4, pp. 663–686, 2020. https://doi.org/10.1108/ijlss-10-2018-0108

[11] P. Badakhshan, B. Wurm, T. Grisold, J. Geyer-Klingeberg, J. Mendling, and J. vom Brocke, "Creating business value with process mining," *Journal of Strategic Information Systems*, vol. 31, no. 4, 2022. https://doi.org/10.1016/j.jsis.2022.101745

[12] M. Martinelli, M. Lippi, and R. Gamberini, "Poka Yoke Meets Deep Learning: A Proof of Concept for an Assembly Line Application," *Applied Sciences*, vol. 12, no. 21, p. 11071, 2022. https://doi.org/10.3390/app122111071

[13] S. Fenner and T. Netland, "Lean service: A contingency perspective," *Operations Management Research*, vol. 16, pp. 1271–1289, 2023. https://doi.org/10.1007/s12063-023-00350-7
[14] J. Van den Bergh, J. Beliën, P. De Bruecker, E. Demeulemeester, and L. De Boeck, "Personnel scheduling: A literature review," *European Journal of Operational Research*, vol. 226, no. 3, pp. 367–385, 2013. https://doi.org/10.1016/j.ejor.2012.11.029
[15] L. Ricciardi Celsi, A. Caliciotti, M. D'Onorio, E. Scocchi, N. A. Sulieman, and M. Villari, "On Predicting Ticket Reopening for Improving Customer Service in 5G Fiber Optic Networks," *Future Internet*, vol. 13, no. 10, p. 259, 2021. https://doi.org/10.3390/fi13100259
[16] S. Agostinelli, F. Chiariello, F. M. Maggi, A. Marrella, and F. Patrizi, "Process mining meets model learning: Discovering deterministic finite state automata from event logs for business process analysis," *Information Systems*, vol. 114, p. 102180, 2023. https://doi.org/10.1016/j.is.2023.102180
[17] V. L. A. Daneida, Aumento del cumplimiento del SLA a través de la Optimización de la Gestión de Servicios (ISO 20000-1) y la Implementación de PDCA en una Empresa de Servicios de TI, 2025. https://repositorioacademico.upc.edu.pe/handle/10757/686974
[18] P. Serirungsun, Y. F. Badir, and B. Frank, "Achieving Digital Transformation: The Effect of Open Innovation Strategies at Different Stages of SME Development in an Emerging Market," *IEEE Transactions on Engineering Management*, 2024. Available: https://doi.org/10.1109/TEM.2024.3487297
[19] H. A. Reijers, "Business Process Management: The evolution of a discipline," *Computers in Industry*, vol. 126, p. 103404, 2021. https://doi.org/10.1016/j.compind.2021.103404
[20] A. R. Permatasari, S. Sulistyo, and P. I. Santosa, "Optimizing IT Services Quality: Implementing ITIL for Enhanced IT Service Management," in *Proceedings of the 11th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, 2024. https://doi.org/10.1109/ICITACEE62763.2024.10762782
[21] S. Poudel, N. Rajkarnikar, D. Shrestha, D. Shrestha, and S. R. Jeong, "Business and Information Technology Strategy Impact on Organizational Performance: A Case Study of Nepal Telecom," in *Lecture Notes in Networks and Systems*, 2023. https://doi.org/10.1007/978-3-031-50158-6_4
[22] S.-C. Haw, K. Ong, L.-J. Chew, K.-W. Ng, and P. Naveen, "Improving the Prediction Resolution Time for Customer Support Ticket System," *Journal of System and Management Sciences*, 2022. https://doi.org/10.33168/JSMS.2022.0601
[23] E. M. Frazzon, L. Stradioto, J. C. Bremen, and P. Onorio, "Data-driven approach for the digital transformation in small and medium-sized enterprises: insights from Brazil," in *Building Digital Technological Capabilities*, 2025. https://doi.org/10.4337/9781803922362.00017
[24] M. F. Van Assen and B. A. Lameijer, "The service concept – a missing link in lean for services," *Total Quality Management & Business Excellence*, vol. 34, no. 15–16, pp. 2027–2045, 2023. Available: https://doi.org/10.1080/14783363.2023.2222064
[25] A. Zangari, M. Marcuzzo, M. Schiavinato, A. Gasparetto, and A. Albarelli, "Ticket automation: An insight into current research with applications to multi-level classification scenarios," *Expert Systems with Applications*, vol. 225, p. 119984, 2023. Available: https://doi.org/10.1016/j.eswa.2023.119984
[26] M. N. İnce and Ç. Taşdemir, "Forecasting Retail Sales for Furniture and Furnishing Items through the Employment of Multiple Linear Regression and Holt–Winters Models," *Systems*, vol. 12, no. 6, p. 219, 2024. Available: https://doi.org/10.3390/systems12060219
[27] J. Macias-Aguayo, L. Garcia-Castro, K. F. Barcia, D. McFarlane, and J. Abad-Moran, "Industry 4.0 and Lean Six Sigma Integration: A Systematic Review of Barriers and Enablers," *Applied Sciences*, vol. 12, no. 22, p. 11321, 2022. Available: https://doi.org/10.3390/app122211321
[28] F. Al-Hawari and H. Barham, "A machine learning based help desk system for IT service management," *Journal of King Saud University – Computer and Information Sciences*, vol. 33, no. 1, pp. 702–718, 2019. https://doi.org/10.1016/j.jksuci.2019.04.001
[29] R. Masoud and S. Basahel, "The Effects of Digital Transformation on Firm Performance: The Role of Customer Experience and IT Innovation," *Digital*, vol. 3, no. 2, 2023. https://doi.org/10.3390/digital3020008
[30] I. L. Sharon Christa and V. Suma, "Predictive analytics in IT Service Management (ITSM)," in *Data Mining and Machine Learning Applications*, 2022. https://doi.org/10.1002/9781119792529.ch7