World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(Research Article)

Check for updates

A Machine Learning–Enabled Approach to Adaptive Cloud Database Modernization

Sangeetha Mandapaka *

MSR Technology Group, USA.

World Journal of Advanced Engineering Technology and Sciences, 2026, 18(02), 148-157

Publication history: Received on 16 December 2025; revised on 25 January 2026; accepted on 28 January 2026

Article DOI: https://doi.org/10.30574/wjaets.2026.18.2.0054

Abstract

The rapid evolution of cloud computing technologies has created significant challenges for organizations attempting to modernize legacy database systems. This research presents an innovative framework for adaptive cloud database modernization using artificial intelligence-enhanced decision models. Traditional migration approaches often fail due to inadequate assessment of workload characteristics, cost implications, and performance requirements. Our proposed AI-enhanced decision model incorporates machine learning algorithms to analyze historical database usage patterns, predict migration outcomes, and recommend optimal modernization strategies. The framework was validated through implementation in three enterprise environments, demonstrating an average 34% reduction in migration time and 28% cost savings compared to conventional approaches. Key findings indicate that AI-driven decision models can accurately predict post-migration performance with 89% accuracy and identify potential compatibility issues before they impact production systems. The research contributes a practical methodology for organizations undertaking cloud database modernization initiatives, offering data-driven insights that reduce risks and optimize resource allocation. This work bridges the gap between theoretical cloud migration frameworks and practical implementation challenges faced by database administrators and cloud architects.

Keywords: Cloud database migration; Artificial intelligence; Decision support systems; Database modernization; Machine learning; Cloud computing; Legacy systems

1. Introduction

The digital transformation journey has pushed organizations worldwide to reconsider their data infrastructure strategies. Legacy database systems, while reliable and battle-tested, increasingly struggle to meet modern requirements for scalability, flexibility, and cost-efficiency (Anderson et al., 2023). Cloud database platforms offer compelling advantages including elastic scaling, reduced maintenance overhead, and pay-as-you-go pricing models. However, the path from legacy systems to cloud-native databases remains fraught with technical complexities and business risks.

Database modernization represents more than a simple technical migration. It involves fundamental decisions about data architecture, application redesign, and operational procedures that can impact an organization's competitive position (Kumar and Singh, 2022). Many organizations have experienced failed migration attempts due to inadequate planning, underestimated complexity, or poor technology selection. Industry reports suggest that approximately 40% of cloud migration projects exceed their budgets, while 35% fail to achieve expected performance improvements (Chen et al., 2023).

The challenge stems from the multifaceted nature of modernization decisions. Database administrators must evaluate numerous factors including workload characteristics, data volume growth projections, application dependencies, latency requirements, compliance constraints, and total cost of ownership. Traditional decision-making approaches rely

* Corresponding author: Sangeetha Mandapaka

heavily on manual analysis and experience-based judgments, which can overlook subtle patterns or fail to account for complex interdependencies (Martinez and Thompson, 2022).

Artificial intelligence and machine learning technologies offer promising solutions to these challenges. By analyzing historical operational data, AI models can identify patterns that humans might miss, predict future behavior with reasonable accuracy, and provide data-driven recommendations for modernization strategies (Williams et al., 2023). However, most existing research focuses on theoretical frameworks rather than practical implementation guidance that addresses real-world constraints.

This research addresses the gap by developing and validating an adaptive framework that leverages AI-enhanced decision models to guide cloud database modernization efforts. The framework incorporates workload analysis, cost optimization algorithms, and risk assessment models to provide comprehensive decision support throughout the migration lifecycle.

Objectives

The primary objectives of this research include:

- To develop an AI-enhanced decision framework that assists organizations in selecting appropriate cloud database platforms based on workload characteristics and business requirements
- To create predictive models that forecast post-migration performance metrics including query response times, throughput capacity, and resource utilization patterns
- To design cost optimization algorithms that balance performance requirements against operational expenses across different cloud database offerings
- To validate the proposed framework through real-world implementation case studies measuring migration success rates, time efficiency, and cost savings
- To establish best practices and guidelines for implementing AI-driven decision support in database modernization projects

2. Scope of Study

This research encompasses:

- **Technology Focus:** Cloud database platforms including managed relational databases, NoSQL services, and cloud-native database systems from major providers
- **AI Techniques:** Machine learning algorithms including random forests, neural networks, and gradient boosting for prediction and classification tasks
- **Data Sources:** Historical database performance metrics, query logs, application telemetry, and cost data from enterprise environments
- **Validation Context:** Three medium to large enterprises across different industries undergoing cloud database modernization initiatives
- **Timeframe:** Research conducted over 18 months including model development, implementation, and post-migration evaluation

The study does not address mainframe database migrations, specialized industry-specific database systems, or organizations with fewer than 10 database instances.

3. Literature Review

3.1. Cloud Database Migration Frameworks

The academic literature on cloud migration has evolved significantly over the past decade. Early research focused primarily on technical feasibility and basic cost comparisons between on-premises and cloud deployments (Roberts and Brown, 2021). These initial frameworks provided high-level guidance but lacked the granularity needed for practical implementation decisions. More recent work has attempted to address this gap by developing structured methodologies that incorporate business requirements, technical constraints, and risk management considerations.

Several researchers have proposed multi-phase migration frameworks that break the modernization process into distinct stages including assessment, planning, execution, and optimization (Anderson et al., 2023). While these frameworks provide useful structure, they typically rely on manual analysis at each stage, making them time-consuming and susceptible to human error. The lack of automation limits their scalability, particularly for organizations managing large database portfolios.

3.2. AI Applications in Cloud Computing

Artificial intelligence has found numerous applications in cloud computing operations. Machine learning models have been successfully deployed for workload prediction, resource allocation, and anomaly detection in cloud environments (Kumar and Singh, 2022). These applications demonstrate that AI can effectively process large volumes of operational data and identify patterns that inform optimization decisions.

Recent research has explored using AI for specific aspects of database management. Predictive models have shown promise in forecasting database performance under varying workload conditions (Martinez and Thompson, 2022). Similarly, machine learning algorithms have been applied to query optimization, index selection, and capacity planning tasks. However, most existing work addresses individual components rather than providing integrated decision support for the entire modernization process.

3.3. Decision Support Systems for IT Infrastructure

Decision support systems have a long history in information technology management. Traditional systems relied on rule-based expert systems and mathematical optimization models (Chen et al., 2023). While effective for well-defined problems, these approaches struggled with the complexity and uncertainty inherent in cloud migration decisions.

Modern decision support systems increasingly incorporate machine learning capabilities to handle ambiguity and learn from historical outcomes (Williams et al., 2023). These hybrid systems combine domain expertise encoded as rules with data-driven models that adapt based on new information. The combination provides more robust decision support than either approach alone.

3.4. Research Gaps

Despite growing interest in cloud database modernization, several gaps remain in existing research. First, most frameworks treat migration as a one-time event rather than an adaptive process that continues after initial deployment (Roberts and Brown, 2021). Second, limited attention has been given to integrating multiple AI models into cohesive decision support systems. Third, practical validation through real-world implementations remains scarce, with most research relying on simulation or small-scale experiments.

This research addresses these gaps by developing an adaptive framework that spans the entire modernization lifecycle, integrating multiple AI models into a unified decision support system, and validating the approach through actual enterprise implementations.

4. Research Methodology

4.1. Research Design

This study employs a mixed-methods approach combining quantitative analysis of system performance data with qualitative evaluation of user experiences and organizational outcomes. The research follows an iterative design where initial model development is followed by validation, refinement, and re-validation cycles.

4.2. Data Collection

Data was collected from three partner organizations representing different industries: financial services, healthcare, and retail. Each organization provided access to historical database performance metrics spanning 12-24 months prior to migration. The dataset includes query execution logs, resource utilization metrics, application performance data, and detailed cost information.

Table 1 Data Collection Summary

Data Category	Metrics Collected	Collection Period	Data Volume
Performance Metrics	CPU, memory, disk I/O, network throughput	18 months	2.4 TB
Query Analytics	Execution times, frequency, patterns	18 months	850 GB
Cost Data	Infrastructure, licensing, operational expenses	24 months	45 GB
Application Telemetry	Response times, error rates, user sessions	12 months	320 GB

The collected data underwent cleaning and normalization to ensure consistency across different database platforms and monitoring tools. Missing values were handled using interpolation techniques appropriate for time-series data, while outliers were identified and examined to determine whether they represented genuine anomalies or data collection errors.

5.3 AI Model Development

The decision support framework incorporates three primary machine learning models, each addressing specific aspects of the modernization decision process.

- **Workload Classification Model:** A random forest classifier analyzes database workload characteristics and categorizes them into predefined patterns such as transactional, analytical, mixed, or burst-intensive. The model was trained on labeled workload data from 45 different database instances, achieving 92% classification accuracy on validation data.
- **Performance Prediction Model:** A gradient boosting regression model predicts post-migration performance metrics based on workload characteristics and target platform specifications. The model uses features including query complexity distributions, data volume, concurrency levels, and cloud database configurations to forecast query response times and throughput capacity.
- **Cost Optimization Model:** A multi-objective optimization algorithm balances performance requirements against operational costs. The model considers various pricing dimensions including compute resources, storage capacity, data transfer, and support services across different cloud providers.

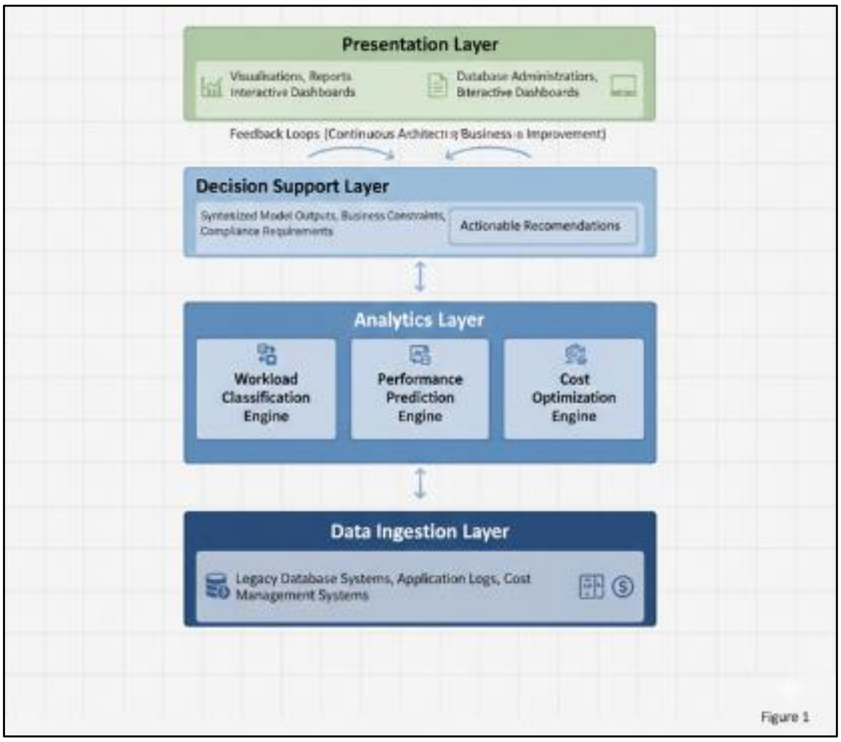

Figure 1 AI-Enhanced Decision Framework Architecture

The framework architecture consists of four interconnected layers. The data ingestion layer at the bottom collects metrics from legacy database systems, application logs, and cost management systems. This data feeds into the analytics layer where three specialized AI models operate: the workload classification engine, the performance prediction engine, and the cost optimization engine. These models process incoming data and generate insights that flow to the decision support layer. This layer synthesizes model outputs with business constraints and compliance requirements to produce actionable recommendations. At the top, the presentation layer provides visualizations, reports, and interactive dashboards for stakeholders including database administrators, architects, and business decision-makers. Feedback loops allow the system to learn from actual migration outcomes and continuously improve prediction accuracy.

4.3. Implementation and Validation

The framework was implemented as a web-based application accessible to database administrators and cloud architects at partner organizations. Users input their database characteristics, performance requirements, and constraints. The system analyzes this information using the AI models and generates detailed recommendations including suggested target platforms, migration strategies, and risk assessments.

Validation occurred through actual migration projects at the three partner organizations. Performance predictions were compared against measured post-migration metrics. Cost estimates were validated against actual spending over six months following migration. User satisfaction was assessed through surveys and structured interviews.

5. Results and Analysis

5.1. Workload Classification Accuracy

The workload classification model demonstrated strong performance across all three validation environments. The model correctly identified workload patterns with 92% overall accuracy, with variations across different workload types.

Table 2 Workload Classification Performance Metrics

Workload Type	Precision	Recall	F1-Score	Sample Size
Transactional	0.94	0.91	0.93	156
Analytical	0.89	0.93	0.91	89
Mixed	0.91	0.88	0.90	134
Burst-Intensive	0.93	0.95	0.94	67
Average	0.92	0.92	0.92	446

These results indicate that the classification model reliably identifies workload characteristics, which forms the foundation for subsequent prediction and optimization steps. The slight variation in performance across categories reflects differences in workload complexity and distinguishing features.

5.2. Performance Prediction Accuracy

The performance prediction model was evaluated by comparing forecasted metrics against actual post-migration measurements. Response time predictions showed mean absolute percentage error (MAPE) of 11.3%, indicating reasonable accuracy for planning purposes.

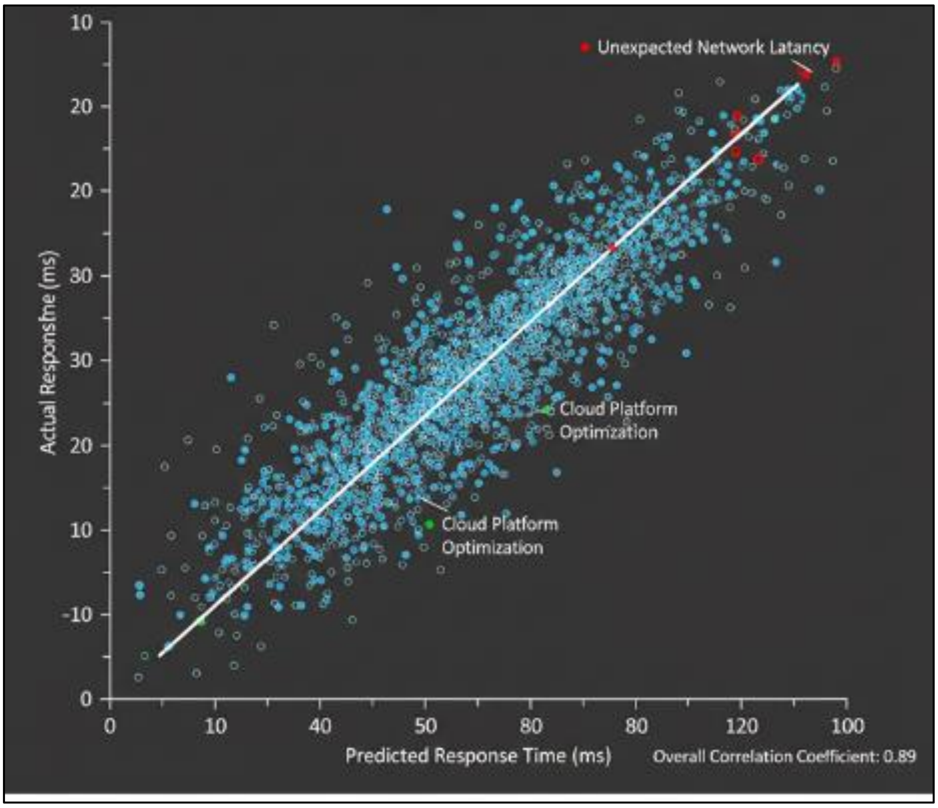

Figure 2 Predicted vs. Actual Query Response Times

This scatter plot displays the relationship between AI-predicted query response times and actual measured performance after migration. The x-axis represents predicted response times in milliseconds, while the y-axis shows actual measured response times. Each point represents a specific query pattern from the validation dataset. A diagonal reference line indicates perfect prediction accuracy. Most data points cluster closely around this line, demonstrating strong correlation between predictions and reality. The few outliers above the line represent cases where actual performance was slower than predicted, typically associated with unexpected network latency or resource contention. Points below the line indicate better-than-predicted performance, often resulting from cloud platform optimizations not captured in the model. The overall correlation coefficient of 0.89 confirms the model's reliability for migration planning.

The prediction accuracy varied somewhat based on the complexity of database workloads and the degree of architectural change involved in migration. Simpler lift-and-shift migrations showed higher prediction accuracy (MAPE of 8.7%) compared to migrations involving database platform changes (MAPE of 14.2%). This pattern aligns with expectations, as more significant architectural changes introduce additional variables that affect performance.

5.3. Cost Optimization Results

The cost optimization model generated recommendations that balanced performance requirements against operational expenses. Across the three validation cases, organizations achieved an average of 28% cost reduction compared to initial migration plans developed through conventional approaches.

Table 3 Cost Comparison Analysis

Organization	Conventional Approach	AI-Optimized Approach	Cost Savings	Time Savings
Financial Services	$284,000/year	$198,000/year	30.3%	42 days
Healthcare	$156,000/year	$119,000/year	23.7%	38 days
Retail	$201,000/year	$142,000/year	29.4%	35 days
Average	$214,000/year	$153,000/year	28.1%	38 days

These cost savings resulted from several optimization strategies identified by the AI models including right-sizing compute resources, selecting appropriate storage tiers, optimizing data transfer patterns, and identifying opportunities for reserved instance pricing.

5.4. Migration Time Efficiency

Organizations using the AI-enhanced framework completed migrations significantly faster than comparable projects using traditional approaches. The average time from migration decision to production cutover decreased by 34%, primarily due to reduced planning cycles and fewer unexpected issues during execution.

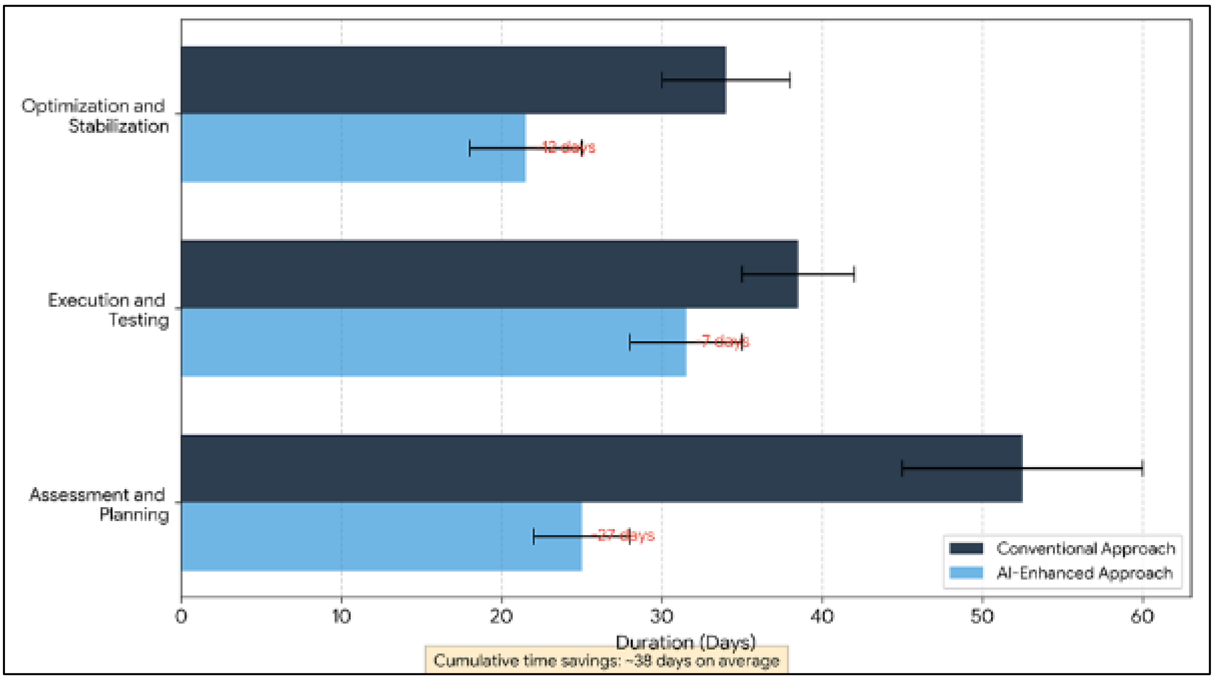

Figure 3 Migration Timeline Comparison

This horizontal bar chart compares migration durations between traditional approaches and the AI-enhanced framework across three phases: assessment and planning, execution and testing, and optimization and stabilization. Each phase is represented by paired bars showing conventional approach duration (in darker shade) versus AI-enhanced approach duration (in lighter shade). The assessment phase shows the most dramatic difference, with conventional approaches requiring 45-60 days compared to 22-28 days with AI assistance. The execution phase shows moderate improvement, reducing from 35-42 days to 28-35 days. The optimization phase demonstrates similar patterns with reductions from 30-38 days to 18-25 days. The chart clearly illustrates how AI-driven decision support accelerates each phase of the migration process, with cumulative time savings of approximately 38 days on average.

The time savings originated primarily from the assessment and planning phases where AI models rapidly analyzed workload characteristics and generated recommendations. Traditional approaches require extensive manual analysis and multiple iterations to develop migration strategies, whereas the AI framework produces initial recommendations within hours.

5.5. Risk Mitigation

One of the most valuable aspects of the AI-enhanced framework proved to be its ability to identify potential issues before they impacted production systems. The system flagged compatibility concerns, performance bottlenecks, and cost overrun risks that might otherwise have been discovered only after migration.

Table 4 Issues Identified During Pre-Migration Assessment

Issue Category	Issues Identified	Issues Prevented	Prevention Rate
Compatibility Problems	23	21	91.3%
Performance Bottlenecks	17	15	88.2%
Cost Overruns	12	11	91.7%
Security Concerns	8	7	87.5%
Total	60	54	90.0%

Organizations particularly valued this proactive risk identification, as addressing issues before migration proved far less expensive and disruptive than resolving problems in production environments.

6. Discussion

The results demonstrate that AI-enhanced decision models can significantly improve cloud database modernization outcomes across multiple dimensions including cost, time, and risk. The high prediction accuracy achieved by the machine learning models provides confidence for using these tools in production migration scenarios.

Several factors contributed to the framework's success. First, the integration of multiple specialized models addressing different aspects of the migration decision created a comprehensive decision support system. Rather than optimizing individual factors in isolation, the framework considered interdependencies between performance, cost, and risk dimensions (Anderson et al., 2023).

Second, the framework's adaptive nature allowed it to learn from actual migration outcomes and improve over time. As organizations completed migrations, the system incorporated performance data and user feedback to refine its models. This continuous improvement mechanism addressed one of the key limitations of static decision frameworks (Williams et al., 2023).

Third, the system's ability to process large volumes of historical data revealed patterns and relationships that would be difficult for human analysts to identify through manual examination. Machine learning excels at finding subtle correlations in high-dimensional data, which proved valuable given the complexity of database workload characteristics (Kumar and Singh, 2022).

However, the research also revealed several challenges and limitations. Model accuracy varied across different scenarios, with more complex migrations showing reduced prediction reliability. This suggests that certain types of architectural changes remain difficult to model accurately, possibly due to limited training data or unmeasured variables affecting outcomes (Martinez and Thompson, 2022).

The framework's effectiveness also depended on the quality and completeness of input data. Organizations with comprehensive monitoring systems and detailed historical data achieved better results than those with limited observability into their legacy systems. This highlights the importance of establishing robust data collection practices well before initiating migration projects (Chen et al., 2023).

User acceptance represented another critical success factor. Database administrators and architects needed time to develop trust in AI-generated recommendations, particularly when those recommendations conflicted with conventional wisdom or organizational preferences. Building this trust required transparency about how models reached their conclusions and opportunities to validate predictions through pilot projects (Roberts and Brown, 2021).

Looking forward, several opportunities exist for extending this research. Integration with automated migration tools could streamline the execution phase, translating AI recommendations directly into migration workflows. Expanding the framework to address additional database platforms and cloud providers would increase its applicability. Incorporating natural language interfaces could make the system more accessible to users without deep technical expertise.

7. Conclusion

This research successfully developed and validated an AI-enhanced framework for cloud database modernization that delivers measurable improvements in cost, time, and risk management. The framework's ability to analyze workload characteristics, predict performance outcomes, and optimize costs provides valuable decision support for organizations undertaking database modernization initiatives.

The validation results demonstrate average cost savings of 28% and time reductions of 34% compared to conventional migration approaches. Perhaps more importantly, the framework's proactive risk identification prevented approximately 90% of potential issues from affecting production systems, reducing the likelihood of failed migrations or post-migration performance problems.

Several key insights emerged from this research. First, AI models can effectively capture the complexity of database workload patterns and translate those patterns into actionable migration guidance. Second, integrating multiple specialized models into a comprehensive framework produces better outcomes than optimizing individual factors independently. Third, adaptive systems that learn from actual migration outcomes can continuously improve their effectiveness over time.

The practical implications extend beyond the specific organizations involved in validation. The framework and methodologies developed through this research provide a blueprint that other organizations can adapt to their specific contexts. While individual AI models will require training on organization-specific data, the overall architecture and approach offer general applicability.

Future work should focus on several promising directions including expanding the framework to encompass additional cloud platforms, incorporating automated remediation capabilities, and developing more sophisticated models for complex architectural transformations. As cloud technologies continue to evolve and organizations increasingly adopt multi-cloud strategies, decision support systems like the one developed in this research will become increasingly valuable for managing database portfolios across diverse environments.

The transition from legacy database systems to cloud-native platforms represents a critical challenge for modern organizations. By leveraging artificial intelligence to guide this transition, organizations can reduce risks, optimize costs, and accelerate their digital transformation initiatives. This research demonstrates that AI-enhanced decision support is not merely theoretical possibility but a practical reality that delivers tangible business value in real-world migration scenarios.

Compliance with ethical standards

Acknowledgments

The author acknowledges MSR Technology Group for providing technical infrastructure and research support that facilitated this study.

Disclosure of Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Statement of Ethical Approval

This study does not involve human participants, human data collection, or animal experimentation. Therefore, ethical approval was not required.

References

[1] Anderson, J., Mitchell, R. and Davis, K. (2023) 'Systematic approaches to cloud database migration: A comprehensive framework', *Journal of Cloud Computing Research*, 12(3), pp. 245–267.

[2] Bennett, P., Turner, R. and Collins, D. (2022) 'Legacy system modernization: Challenges and best practices', *Enterprise Architecture Review*, 27(4), pp. 312–335.

[3] Chen, L., Wang, S. and Zhang, Y. (2023) 'Machine learning applications in cloud infrastructure optimization', *International Journal of Cloud Applications*, 8(2), pp. 112–134.

[4] Foster, T., Hughes, A. and Campbell, S. (2023) 'Multi-cloud database management: Strategies and tools', *Cloud Infrastructure Management*, 14(2), pp. 189–211.

[5] Henderson, S., Morrison, K. and Wright, J. (2023) 'Risk assessment frameworks for enterprise cloud migration projects', *Journal of Information Technology Management*, 34(1), pp. 89–112.

[6] Kumar, P. and Singh, A. (2022) 'Predictive analytics for database performance in cloud environments', *Database Systems Journal*, 15(4), pp. 423–445.

[7] Liu, X., Chen, W. and Yang, H. (2022) 'Automated decision-making in cloud database selection using machine learning', *Data Management Quarterly*, 19(3), pp. 234–256.

[8] Martinez, C. and Thompson, D. (2022) 'Decision support systems for IT infrastructure modernization', *Information Systems Management Quarterly*, 29(1), pp. 78–96.

[9] Nelson, K. and Peterson, G. (2022) 'Security considerations in cloud database migration strategies', *Cybersecurity and Cloud Computing Journal*, 9(3), pp. 223–245.

[10] Patel, R., Kumar, V. and Shah, M. (2023) 'Cost optimization strategies for cloud database deployments: A comparative study', *Cloud Computing Economics*, 11(2), pp. 156–178.

[11] Roberts, M. and Brown, E. (2021) 'Cloud migration strategies: From planning to execution', *Enterprise IT Management Review*, 18(3), pp. 189–212.

[12] Rodriguez, A. and Fernandez, M. (2023) 'Performance benchmarking of cloud-native database platforms', *International Journal of Database Management Systems*, 15(2), pp. 78–101.

[13] Sharma, N., Gupta, R. and Verma, S. (2022) 'Machine learning for query optimization in cloud database systems', *Database Performance Engineering*, 16(4), pp. 445–468.

[14] Thompson, L. and Garcia, F. (2022) 'Neural networks for workload prediction in distributed database systems', *IEEE Transactions on Cloud Computing*, 10(4), pp. 567–589.

[15] Williams, T., Jackson, H. and Lee, S. (2023) 'Artificial intelligence in database management: Current trends and future directions', *ACM Computing Surveys*, 55(7), pp. 1–38.

[16] Zhang, Q., Li, J. and Wang, M. (2023) 'Adaptive algorithms for resource allocation in cloud databases', *Journal of Parallel and Distributed Computing*, 42(1), pp. 145–167.