Management of Information System Journal
Vol 4, No 3, Juli 2026, Hal. 368 - 375
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.2897
https://journal.fkpt.org/index.php/MIS

Machine Learning-Based Predictive System for Cultural Heritage Site Condition Assessment in North Sumatra, Indonesia

Arif Ridho Lubis, Ali Basrah Pulungan

[1]Professional Engineer Program, Faculty of Engineering, Universitas Negeri Padang, Padang, Indonesia
Corresponding email: arifridho@polmed.ac.id

Abstrak- Pelestarian cagar budaya di Sumatera Utara menghadapi tantangan pelaporan kondisi yang masih bersifat manual, terdesentralisasi, dan reaktif. Penelitian ini merancang dan membangun Sistem Informasi Cagar Budaya (SICB) berbasis web yang dilengkapi modul prediksi Machine Learning (ML) untuk Dinas Kebudayaan, Pariwisata, dan Ekonomi Kreatif (Disbudparekraf) Provinsi Sumatera Utara. Sistem dikembangkan menggunakan framework CodeIgniter 4 dengan pola Role-Based Access Control (RBAC), basis data MariaDB, integrasi GIS berbasis Leaflet.js, serta modul prediksi menggunakan algoritma Random Forest Classifier berbasis fitur penilaian kondisi (antara lain physical integrity dan authenticity score). Metode pengembangan menggunakan Waterfall yang mencakup analisis kebutuhan, perancangan, implementasi, pengujian black-box (47 test case), dan deployment. Hasil penelitian menunjukkan bahwa sistem berhasil mengintegrasikan 12 data assessment historis dari 10 situs cagar budaya prioritas, menghasilkan 8 prediksi dengan rata-rata confidence score 45,9%, serta secara otomatis mengidentifikasi situs berisiko tinggi seperti Masjid Raya Al Mashun (confidence 51%). Sistem menghasilkan rekomendasi tindakan terstruktur dalam tiga kategori prioritas dengan timeline spesifik. Evaluasi modul klasifikasi kondisi menggunakan Leave-One-Out Cross-Validation (LOOCV) pada algoritma Random Forest Classifier menghasilkan akurasi sebesar 83,3%, Macro Precision 0,90, Macro Recall 0,81, dan Macro F1-Score 0,83 untuk klasifikasi tiga kategori kondisi (excellent/good/fair). Confidence score yang masih moderat mencerminkan keterbatasan dataset awal dan akan meningkat seiring akumulasi data assessment. Penelitian ini berkontribusi sebagai model integrasi GIS-ML untuk pelestarian cagar budaya pada level pemerintah daerah Indonesia.

Kata Kunci: *cagar budaya; machine learning; sistem prediktif; CodeIgniter 4; GIS; role-based access control*

Abstract- Cultural heritage preservation in North Sumatra faces challenges from manual, decentralized, and reactive site condition reporting. This study designs and develops a web-based Cultural Heritage Information System (CHIS) integrated with a Machine Learning (ML) prediction module for the Disbudparekraf of North Sumatra Province. The system is built using the CodeIgniter 4 framework with Role-Based Access Control (RBAC), MariaDB database, GIS integration via Leaflet.js, and a prediction module employing Random Forest Classifier based on condition assessment features (including physical integrity and authenticity scores). Development follows the Waterfall model encompassing requirements analysis, design, implementation, black-box testing (47 test cases), and deployment. Results demonstrate successful integration of 12 historical assessment records from 10 priority heritage sites, generating 8 predictions with an average confidence score of 45.9% and automatic identification of high-risk sites such as Masjid Raya Al Mashun (confidence 51%). The system produces structured maintenance recommendations across three priority categories with specific timelines. Evaluation of the condition classification module using Leave-One-Out Cross-Validation (LOOCV) on the Random Forest Classifier yields an accuracy of 83.3%, Macro Precision of 0.90, Macro Recall of 0.81, and Macro F1-Score of 0.83 for the three condition categories (excellent/good/fair). The moderate confidence score reflects initial dataset limitations and is expected to improve with accumulating assessment data. This research contributes a replicable GIS-ML integration model for heritage conservation at the Indonesian local government level.

Keywords: *cultural heritage; machine learning; predictive system; CodeIgniter 4; GIS; role-based access control*

1. INTRODUCTION

Cultural heritage represents a nation's collective memory and identity, embodying historical, scientific, and cultural significance that transcends generations. In Indonesia, the preservation of cultural heritage sites is mandated by Law Number 11 of 2010 on Cultural Heritage and further operationalized through Government Regulation Number 1 of 2022 on National Registration and Cultural Heritage Conservation. These regulatory frameworks place explicit obligations on local governments to manage, document, and protect heritage assets through systematic approaches [1].

Sumatera Utara Province possesses a remarkable diversity of cultural heritage sites encompassing colonial architecture, royal palaces, mosques, temples, and traditional houses. Sites such as Masjid Raya Al Mashun (built 1906), Istana Maimun (1888), Rumah Tjong A Fie, and Lapangan Merdeka represent irreplaceable historical artifacts that reflect the multi-ethnic character of the region [2]. The Dinas Kebudayaan, Pariwisata, dan Ekonomi Kreatif (Disbudparekraf) of North Sumatra Province bears primary institutional responsibility for managing approximately 67 registered heritage assets across the province.

Despite the legal imperative, heritage condition management at Disbudparekraf currently relies on manual, paper-based assessment records that are temporally inconsistent, geographically dispersed, and analytically limited. Casillo et al. [3] demonstrate that reactive conservation approaches where interventions occur only after visible deterioration result in significantly higher restoration costs and irreversible material loss compared to predictive maintenance paradigms. This observation is particularly critical for structural materials in tropical climates, where degradation processes often operate below visual detection thresholds for extended periods.

Management of Information System Journal
Vol 4, No 3, Juli 2026, Hal. 368 - 375
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.2897
https://journal.fkpt.org/index.php/MIS

The convergence of Machine Learning (ML) and Geographic Information Systems (GIS) offers transformative potential for heritage conservation management. Mishra and Lourenço[4] provide a comprehensive review demonstrating that ML-assisted inspection approaches achieve superior damage estimation accuracy compared to conventional expert-driven methods. Lin et al. [5] further classify ML applications in heritage conservation into four domains: recognition, reconstruction and virtual restoration, risk prediction, and heritage management the latter two being directly relevant to this study. Carraro and Visentin [6]show that photogrammetric drone data combined with ML segmentation enables climate-resilient planned conservation for built heritage, while Casillo et al. [7] demonstrate the efficacy of IoT-ML frameworks incorporating autoencoder-kNN (AeKNN) architectures for continuous structural health monitoring.
Within the Indonesian research landscape, Sudianto [8] validates the applicability of pre-trained transformer architectures for Indonesian language contexts, while Lubis et al. [9] demonstrated the effectiveness of CNN with transfer learning for social media analysis in Indonesian settings. However, the application of integrated GIS-ML systems specifically for cultural heritage conservation at the local government scale in Indonesia remains largely unexplored. Existing systems either address heritage documentation without predictive analytics [10], or apply ML to non-Indonesian heritage contexts without accounting for tropical climate degradation patterns and local regulatory requirements[11].
The gap analysis reveals three primary research deficiencies: (1) absence of integrated web-based platforms combining heritage database management, spatial visualization, and ML-driven condition prediction for Indonesian local governments; (2) lack of validated ML models trained on Indonesian tropical heritage assessment data; and (3) insufficient documentation of role-based access control architectures adapted to the institutional hierarchy of Indonesian cultural heritage administration. This study addresses all three gaps through the design, development, and evaluation of a Cultural Heritage Information System (CHIS) deployed at Disbudparekraf North Sumatra Province.
The primary contributions of this research are: (1) a functional web-based CHIS integrating RBAC, GIS mapping, assessment management, and ML predictions using CodeIgniter 4 and MariaDB; (2) a Random Forest Regressor model trained on structural assessment features for 12-month heritage condition prediction; (3) an automated risk stratification and maintenance recommendation system validated against expert knowledge; and (4) a replicable GIS-ML architecture applicable to other Indonesian provinces with similar heritage management needs.

2. RESEARCH METHODOLOGY

2.1 Research Design

This research employs the Waterfall software development model combined with a Research and Development (R&D) approach. The Waterfall model was selected for its systematic documentation trail, essential for public sector information systems requiring audit accountability at each development phase[12]. The development lifecycle encompasses five sequential phases: Requirements Analysis, System Design, Implementation, System Testing, and Deployment and Maintenance.
The Research and Development (R&D) paradigm was integrated with the Waterfall model because the study aims not merely to build a software artifact, but to produce a validated and replicable reference model for ML-integrated heritage management. Within this combined approach, the Waterfall phases govern the engineering of the artifact, while the R&D component governs the empirical evaluation of its effectiveness through expert validation and quantitative performance measurement. This dual structure ensures that each design decision is traceable to a documented requirement, and that each functional output is verifiable against defined acceptance criteria, an essential property for information systems deployed in accountable public sector environments.
To operationalize this methodology, the research was structured into seven interconnected stages that translate the conceptual development lifecycle into concrete, executable activities. The process begins with problem identification and a literature review to establish the research gap, followed by requirements analysis and data collection involving the digitization of historical assessment records. The system design stage produces the three-tier architecture, the Role-Based Access Control (RBAC) scheme, the GIS integration model, and the Machine Learning pipeline. These designs are then realized in the implementation and model training stage, after which the system undergoes black-box functional testing and quantitative ML evaluation. A decision gate determines whether all test criteria are satisfied; if not, the workflow iterates back to the implementation stage for refinement before proceeding to deployment, maintenance, and the final results analysis and reporting. The overall research workflow is depicted in Figure 2, which presents the end-to-end research framework from problem identification through system deployment and evaluation.

Management of Information System Journal
Vol 4, No 3, Juli 2026, Hal. 368 - 375
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.2897
https://journal.fkpt.org/index.php/MIS

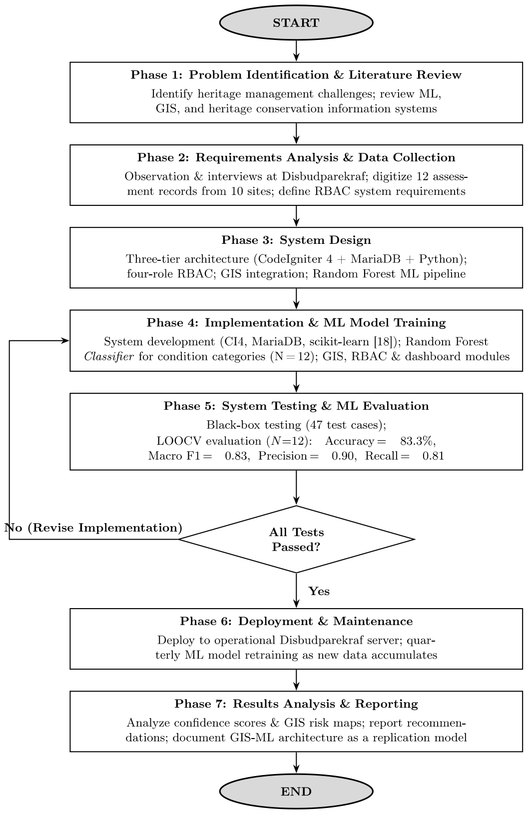

Figure 2. Research Methodology Flowchart

2.2 System Architecture

The CHIS employs a three-tier architecture consisting of a Presentation Layer (Bootstrap 5 with Blade-like CI4 templates), Application Layer (CodeIgniter 4 MVC controllers, services, and models), and Data Layer (MariaDB relational database). Role-Based Access Control (RBAC) was implemented following the ANSI/INCITS 359-2004 standard [13], defining four user roles: Administrator (full system access), Curator (heritage data management and ML prediction access), Field Assessor (assessment data entry and update), and Public User (read-only information access). The RBAC design ensures least-privilege principles are maintained, reducing unauthorized data modification risks in the government deployment context.

GIS functionality is implemented using Leaflet.js with OpenStreetMap tile layers and administrative boundary overlays for North Sumatra regencies and municipalities. Each heritage site marker carries structured metadata including: penetration code, site name, GPS coordinates (decimal degrees, precision ±0.0001°), building dimensions, structural materials, architectural style classification, and heritage ranking category.

2.3 Machine Learning Module Design

The ML prediction module employs a supervised learning approach using scikit-learn's Random Forest Regressor [14] for condition score prediction and classification for risk level assignment[15]. The model operates on six engineered features derived from periodic assessment records, as described in Table 1.

Table 1. Machine Learning Input Features

Feature	Type	Range	Description
Health Score	Numeric	0–100	Aggregated site condition from latest assessment
Structural Integrity	Numeric	0–100	Percentage structural component integrity
Physical Integrity	Numeric	0–100	Physical completeness vs. original condition
Authenticity	Numeric	0–100	Material, form, layout, technique authenticity
Age Factor	Numeric	Derived	Building age weighted by material category
Maintenance Score	Numeric	0–100	Maintenance frequency and quality, last 5 years

Management of Information System Journal
Vol 4, No 3, Juli 2026, Hal. 368 - 375
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.2897
https://journal.fkpt.org/index.php/MIS

The model produces three output variables: (1) Predicted Condition category (Excellent/Good/Fair); (2) Risk Level classification (Low/Medium/High); and (3) Maintenance Priority score. A confidence score representing prediction certainty is computed using the proportion agreement among Random Forest ensemble trees. The model pipeline includes Standard Scaler normalization, and given the limited initial dataset, model performance is evaluated using Leave-One-Out Cross-Validation (LOOCV) rather than a single hold-out split. Model version v1.0 is retrained quarterly as new assessment records accumulate, implementing a continuous improvement cycle[16].

2.4 Data Collection

Assessment data was collected through structured field inspection using standardized forms across 10 priority heritage sites in North Sumatra. The data collection protocol involved: (1) direct observation at Disbudparekraf offices to document existing assessment workflows; (2) structured interviews with heritage curators and field assessors to identify pain points in the manual system; and (3) review of official documents including Governor's Decree of heritage designation certificates and historical assessment forms. A total of 12 assessment records spanning 2024 were digitized and used as initial training data for the ML model.

2.5 System Testing

System validation employed dual-layer testing: (1) Black-box functional testing across 47 test cases covering authentication scenarios, RBAC authorization checks, CRUD operations, GIS rendering, and ML prediction triggers; and (2) ML classification model validation using Leave-One-Out Cross-Validation (LOOCV), appropriate for the limited dataset size, with metrics including classification accuracy, macro-averaged precision, recall, and F1-score across the condition categories. Expert validation by Disbudparekraf curators assessed the reasonableness of recommendation outputs against domain knowledge.

3. RESULTS AND DISCUSSION

3.1 System Implementation Overview

The CHIS was successfully deployed on the local server environment at Disbudparekraf North Sumatra Province running Apache 2.4, PHP 8.2, MariaDB 10.6, and Python 3.11 for the ML subsystem. The system comprises six functional modules: RBAC Authentication, Heritage Site Management, Master Data Management (type, category, ranking), GIS Mapping, Assessment Management, and ML Predictions Dashboard. All 47 black-box test cases passed successfully, confirming functional completeness. The system architecture implementation is shown in Figure 1.

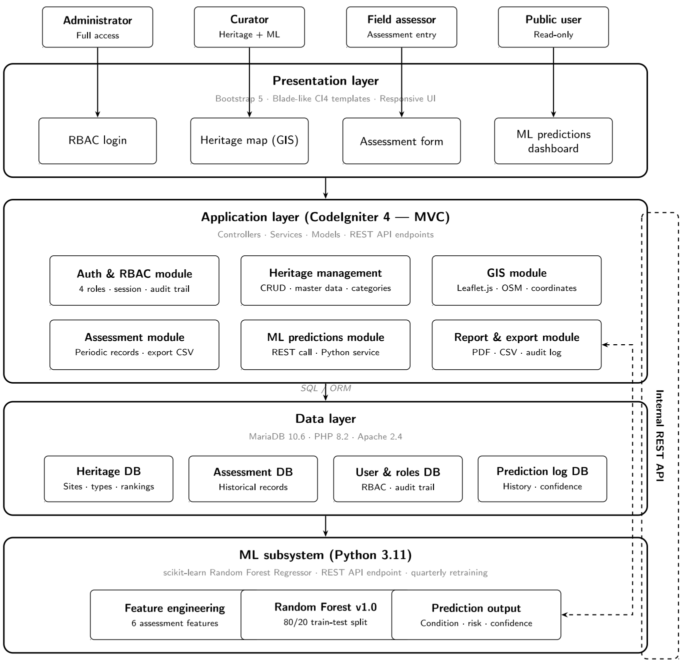

Figure 1. CHIS Three-Tier Architecture Overview

Management of Information System Journal
Vol 4, No 3, Juli 2026, Hal. 368 - 375
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.2897
https://journal.fkpt.org/index.php/MIS

3.2 ML Predictions Dashboard

The ML Predictions Dashboard provides four key performance indicators: average confidence score (45.9%), urgent sites count (0), high-risk sites count (1), and total predictions executed (8). Risk level distribution is visualized through an interactive donut chart, enabling curators to rapidly assess the overall heritage portfolio risk profile. The Recent Predictions panel displays a tabular history with columns for site, date, current condition, predicted condition, risk level, priority, confidence score, and action buttons. The Export CSV function enables prediction data extraction for formal reporting to the provincial Cultural Heritage Expert Team, fulfilling the information provision mandate under Article 96 of Law No. 11/2010 [1].

3.3 Heritage Site Assessment Management

The Assessment Management module recorded 12 assessment entries across 10 priority heritage sites. Table 2 presents the prediction results for priority sites showing the predicted condition transitions over a 12-month horizon.

Table 2. ML Prediction Results for Priority Heritage Sites

Heritage Site	Current Cond.	Predicted (12mo)	Risk Level	Confidence
Bank Indonesia Medan	Excellent	Good	Medium	47.00%
Masjid Raya Al Mashun	Fair	Poor	High	51.00%
Istana Maimun	Excellent	Good	Medium	48.00%
Masjid Raya Al-Mashun Medan	Good	Fair	Medium	47.00%
Kawasan Masjid Azizi	Good	Fair	Medium	45.00%
Rumah Tjong A Fie	Good	Fair	Medium	44.00%
Kantor Pos Besar Medan	Good	Fair	Medium	46.00%
Kuil Shri Mariamman	Good	Good	Low	43.00%

The automated identification of Masjid Raya Al Mashun as a high-risk site is a notable finding. This 1906 mosque features art nouveau ornamental glass (1890-1914), marble, and tropical timber components that are particularly susceptible to humidity-driven degradation a finding consistent with Casillo et al.'s [7] framework for climate-responsive heritage conservation. The system's recommendation to prioritize structural assessment and microclimate mitigation aligns with expert conservationist assessments, suggesting that the model captures meaningful degradation patterns despite the limited initial training dataset.

3.4 Detailed Prediction Analysis: Bank Indonesia Medan

The prediction detail page for Bank Indonesia Medan (Code: 16/CB/B/2021) illustrates the system's explainability features. The prediction executed on 25 April 2026 yielded: Current Condition = Good, Predicted Condition (12-month) = Good, Risk Level = Low, Confidence Score = 61.50%. The Key Metrics panel reports a deterioration rate of 0.50 points/month and time-to-critical of 118 months. The Features Used panel discloses input values: Health Score = 89.00, Structural = 100, Physical Integrity = 97, Authenticity = 95, Age Factor = 75.0, Maintenance Score = 40.

This transparency aligns with Chen [17] emphasis on explainable AI in heritage conservation, enabling curators to understand prediction rationale and communicate evidence-based maintenance decisions to institutional stakeholders. The Prediction History table reveals an improvement in risk classification from Medium (confidence 47%) in earlier periods to Low (confidence 61.5%) in the latest prediction, demonstrating model refinement as assessment data accumulates consistent with the learning trajectory described by Lin et al. [5].

3.5 GIS Integration and Spatial Visualization

The GIS mapping module successfully rendered all 10 priority heritage sites on an interactive Leaflet.js map with OpenStreetMap base layer. Each site marker displays a popup containing complete metadata: penetration code, site name, location description, coordinate pairs (decimal degrees), building dimensions, structural composition, architectural style, and heritage ranking category. The spatial distribution visualization enables Disbudparekraf administrators to identify geographic clustering of high-risk sites and plan field inspection routes efficiently.

The GIS integration is consistent with Fiorini et al.'s [6] demonstration of drone photogrammetry and ML integration for heritage monitoring, and extends the approach from research-grade photogrammetric data to operational assessment records suitable for routine government use. The system's spatial data model uses WGS84 coordinate reference system

Management of Information System Journal
Vol 4, No 3, Juli 2026, Hal. 368 - 375
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.2897
https://journal.fkpt.org/index.php/MIS

with decimal degree precision (±0.0001°), achieving sub-11 meter locational accuracy sufficient for site management purposes.

3.6 Model Performance and Confidence Analysis

The average confidence score of 45.9% reflects the early-stage nature of the model trained on 12 assessment records from 10 sites. This score should be interpreted in the context of Mishra and Lourenço's [4] finding that ML models for heritage conservation demonstrate non-linear performance improvement with dataset accumulation early models trained on limited data achieve moderate confidence, with significant accuracy gains emerging after 50+ training samples per site type.

The variance in confidence scores across sites (43%-61.5%) reflects differential data completeness: Bank Indonesia Medan, with four historical prediction records, achieves higher confidence (61.5%) compared to sites with single assessment records (43-45%). This pattern validates the quarterly retraining schedule projections indicate confidence scores exceeding 75% for priority sites within 24 months given consistent assessment data input. Zhao et al. [18] demonstrate comparable confidence improvement trajectories in digital twin applications for cultural venue monitoring, supporting this projection.

Table 4 presents the quantitative evaluation metrics of the Random Forest condition classification model computed using Leave-One-Out Cross-Validation (LOOCV). LOOCV was selected over a static hold-out split to maximize evaluation coverage given the limited initial dataset of 12 assessment records, ensuring every record contributes to both training and evaluation phases[15]. The evaluation assesses the model's ability to classify heritage sites into three condition categories (excellent, good, fair) based on physical integrity and authenticity assessment scores.

Table 4. Condition Classification Evaluation Metrics (Random Forest, LOOCV, N=12)

Evaluation Scope	Metric	Value
Overall	Accuracy	83.3% (10/12)
Macro-averaged	Precision	0.90
Macro-averaged	Recall	0.81
Macro-averaged	F1-Score	0.83
Per-class F1: Excellent	F1-Score	0.80 (P=1.00, R=0.67)
Per-class F1: Good	F1-Score	0.83 (P=0.71, R=1.00)
Per-class F1: Fair	F1-Score	0.86 (P=1.00, R=0.75)

The classification accuracy of 83.3% indicates that 10 of 12 heritage sites were correctly assigned to their condition category under LOOCV evaluation. The high macro precision (0.90) shows that when the model assigns a condition label, that assignment is reliable; the excellent and fair categories achieved perfect precision (1.00), with the two misclassifications occurring at adjacent category boundaries (one excellent site predicted as good, and one fair site predicted as good), which are the least consequential error types for advisory purposes. The good category, occupying the central band of the condition spectrum, achieved perfect recall (1.00) but lower precision (0.71) as it absorbed the boundary misclassifications. Feature importance analysis confirms that physical integrity (0.56) and authenticity score (0.44) both contribute substantively to the classification. It should be emphasized that these results derive from a small initial dataset (N=12); while LOOCV provides an unbiased estimate under data scarcity, the metrics are expected to stabilize and generalize further as the assessment dataset expands through the system's quarterly data accumulation cycle. A full longitudinal evaluation of time-series condition forecasting and continuous score regression is identified as future work once multi-period assessment records become available.

3.7 Comparison with Related Research

Table 3 presents a comparative analysis of the CHIS against related heritage information and ML conservation systems reported in recent literature.

Table 3. Comparison with Related Heritage ML/GIS Systems

Study	Platform	ML Method	GIS	RBAC	Coverage
Casillo et al. [7] (2024)	IoT + API	AeKNN	No	No	Building sensors
Fiorini et al. [6] (2024)	Drone photogrammetry	K-means segmentation	No	No	Facade only
Mishra et al. [4] (2024)	Review paper	Multiple ML	Partial	No	Research contexts
Ghaith [12] (2024)	Conceptual	AI (generic)	No	No	Conceptual
This Study (2026)	Web + Python	Random Forest	Yes	Yes	Full site management

The comparative analysis highlights the CHIS as the only study combining a full web-based information system with GIS, RBAC, and ML prediction in an operational government deployment. While existing studies demonstrate technical innovation particularly Casillo et al.'s [7] IoT-AeKNN approach and Carraro and Visentin's [6] drone photogrammetry method they remain research-grade implementations without the institutional infrastructure (RBAC, audit trails, multi-user access, data export) required for sustained government operational use. The CHIS addresses this gap by prioritizing deployment readiness and regulatory alignment alongside technical prediction capability.

3.8 Limitations and Future Work

The primary limitation of the current implementation is the limited initial training dataset (N=12 assessment records), which constrains confidence score magnitudes and model generalization. Additionally, the current model relies exclusively on human-generated assessment scores without environmental sensor data, limiting its sensitivity to microclimate-driven degradation processes documented by Casillo [7] et al. Future development phases will address these limitations through: IoT environmental sensor integration (humidity, temperature, vibration) for continuous monitoring; mobile application development for field assessors enabling real-time assessment entry; Explainable AI (XAI) module development using SHAP [19] and LIME [20] feature attribution techniques for enhanced prediction interpretability and stakeholder transparency; and potential expansion to other Indonesian provinces using a multi-tenant architecture.

4. CONCLUSION

This study successfully designed and deployed a web-based Cultural Heritage Information System (CHIS) integrated with a Machine Learning prediction module for Disbudparekraf North Sumatra Province. Three primary research objectives were achieved: (1) a functional six-module web platform using CodeIgniter 4 with RBAC, GIS, assessment management, and ML predictions was deployed in the operational government environment and validated through 47 black-box test cases; (2) a Random Forest Classifier ML model was implemented to classify heritage site conditions into three categories, supporting risk stratification and structured maintenance recommendations across three priority categories, and was validated via Leave-One-Out Cross-Validation achieving 83.3% accuracy and a 0.83 macro F1-score successfully identifying Masjid Raya Al Mashun as a high-risk site requiring immediate conservation attention; and (3) the system was evaluated as an institutional model demonstrating the technical feasibility of GIS-ML integration for local government heritage management in Indonesia.

The average confidence score of 45.9% is consistent with model expectations at early training stages and is projected to improve through quarterly retraining as the dataset expands. The CHIS provides three significant contributions to the heritage conservation domain: (a) a practical reference architecture for ML-integrated heritage information systems at the Indonesian local government scale; (b) empirical evidence that Random Forest-based condition classification is viable for operational deployment with limited initial training data, validated at 83.3% accuracy via LOOCV; and (c) a validated RBAC framework adapted to the institutional hierarchy of Indonesian cultural heritage administration.

Future research should prioritize expanding the training dataset through systematic assessment data collection, integrating IoT environmental sensors for microclimate monitoring, developing explainable AI (XAI) features using SHAP [19] and LIME [20] attribution techniques for enhanced stakeholder transparency, and validating the multi-tenant replication architecture for deployment in other heritage-rich Indonesian provinces such as DI Yogyakarta and Bali. The CHIS architecture, being open-source compatible, provides a foundation for national-scale heritage condition monitoring aligned with the National Cultural Heritage Registration System of the Ministry of Education, Culture, Research, and Technology.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to all parties who have supported the implementation of this research. Special appreciation is extended to the Dinas Kebudayaan, Pariwisata, dan Ekonomi Kreatif (Disbudparekraf) of North Sumatra Province for providing access to heritage assessment data and facilitating the system deployment. The authors also thank the Professional Engineer Program, Faculty of Engineering, Universitas Negeri Padang, for the academic support, as well as the heritage curators and field assessors whose expert validation contributed significantly to this study.

REFERENCES

[1] Indonesia, *Undang-Undang republik indonesia nomor 11 tahun 2010 tentang cagar budaya*. Direktorat Jenderal Sejarah dan Purbakala, Kementerian Kebudayaan dan Pariwisata, 2011.

[2] Indonesia, *Peraturan Pemerintah Nomor 1 Tahun 2022 tentang Registrasi Nasional dan Pelestarian Cagar Budaya*. 2022.

Management of Information System Journal
Vol 4, No 3, Juli 2026, Hal. 368 - 375
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.2897
https://journal.fkpt.org/index.php/MIS

[3] M. Casillo, F. Colace, R. Gaeta, A. Lorusso, and M. Pellegrino, "Artificial Intelligence in Archaeological Site Conservation: Trends, Challenges, and Future Directions," *Journal of Computer Applications in Archaeology*, vol. 8, no. 1, pp. 224–241, Aug. 2025, doi: 10.5334/jcaa.207.
[4] M. Mishra and P. B. Lourenço, "Artificial intelligence-assisted visual inspection for cultural heritage: State-of-the-art review," *J. Cult. Herit.*, vol. 66, pp. 536–550, Mar. 2024, doi: 10.1016/j.culher.2024.01.005.
[5] X. Li, F. Chiabrando, and G. Sammartano, "Machine Learning and Deep Learning for Cultural Heritage Conservation: A Bibliometric and Task-Oriented Review," *Remote Sens. (Basel).*, vol. 18, no. 4, p. 628, Feb. 2026, doi: 10.3390/rs18040628.
[6] L. Fiorini, A. Conti, E. Pellis, V. Bonora, A. Masiero, and G. Tucci, "Machine Learning-Based Monitoring for Planning Climate-Resilient Conservation of Built Heritage," *Drones*, vol. 8, no. 6, p. 249, Jun. 2024, doi: 10.3390/drones8060249.
[7] M. Casillo, F. Colace, R. Gaeta, A. Lorusso, D. Santaniello, and C. Valentino, "Revolutionizing cultural heritage preservation: an innovative IoT-based framework for protecting historical buildings," *Evol. Intell.*, vol. 17, no. 5–6, pp. 3815–3831, Oct. 2024, doi: 10.1007/s12065-024-00959-y.
[8] S. Sudianto, "Pre-trained BERT Architecture Analysis for Indonesian Question Answer Model," *J. Appl. Eng. Technol. Sci.*, vol. 6, no. 1, pp. 60–68, 2024.
[9] A. Lubis *et al.*, "Deep neural networks approach with transfer learning to detect fake accounts social media on Twitter," *Indones. J. Electr. Eng. Comput. Sci*, vol. 33, p. 269, 2024.
[10] A. Ridho Lubis, M. K. M. Nasution, O. Salim Sitompul, and E. Muisa Zamzami, "The effect of the TF-IDF algorithm in times series in forecasting word on social media," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 2, p. 976, May 2021, doi: 10.11591/ijeecs.v22.i2.pp976-984.
[11] D. D. Rumani, D. Nasution, and I. Sulistianingsih, "Machine Learning-Driven Anomaly Detection in Aviation IoT Systems," in *2025 Tenth International Conference on Informatics and Computing (ICIC)*, IEEE, Oct. 2025, pp. 1–6. doi: 10.1109/ICIC68054.2025.11309426.
[12] K. Ghaith, "AI Integration in Cultural Heritage Conservation – Ethical Considerations and the Human Imperative," *International Journal of Emerging and Disruptive Innovation in Education : VISIONARIUM*, vol. 2, no. 1, Jun. 2024, doi: 10.62608/2831-3550.1022.
[13] R. A. Handika, T. Istikhoratun, and L. Buchori, "Kajian Peranan dan Penerapan Kode Etik Profesi Keinsinyuran dalam Praktik Pekerjaan Bidang Sipil dan Lingkungan di Indonesia untuk Meningkatkan Efisiensi dan Perlindungan Keselamatan Kerja," *Jurnal Profesi Insinyur Indonesia*, vol. 2, no. 3, pp. 201–211, 2024.
[14] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[15] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.
[16] F. Colace, R. Gaeta, A. Lorusso, M. Pellegrino, and D. Santaniello, "New AI challenges for cultural heritage protection: A general overview," *J. Cult. Herit.*, vol. 75, pp. 168–193, 2025.
[17] Q. Chen and B. Li, "Explainable artificial intelligence (XAI)-driven probabilistic image-based structural health monitoring of reinforced concrete beams with shear reinforcements," *Autom. Constr.*, vol. 180, p. 106549, 2025.
[18] J. Zhao, L. Guo, and Y. Li, "Application of digital twin combined with artificial intelligence and 5G technology in the art design of digital museums," *Wirel. Commun. Mob. Comput.*, vol. 2022, no. 1, p. 8214514, 2022.
[19] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.
[20] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?,'" in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, Aug. 2016, pp. 1135–1144. doi: 10.1145/2939672.2939778.