International Journal of Science, Strategic Management and Technology
Volume **02** Issue **03** March-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

An Intelligent and Secure Framework for Land Record Digitization using Machine Learning

Author Details:

Kishan Bhatacharjee[1], Dibakar Saha[2]

[1] Computer Science and Engineering, ICFAI University, Agartala, Tripura,India
[2] Software World, Agartala,Tripura,India

Corresponding Author Email: kishanweb2012@email.com | ORCID: https://orcid.org/ 0009-0002-9149-7923

https://doi.org/10.55041/ijsmt.v2i3.011

Cite this Article: Bhatacharjee, D. S. ,. K. (2026). An Intelligent and Secure Framework for Land Record Digitization using Machine Learning. International Journal of Science, Strategic Management and Technology, *02*(03). https://doi.org/10.55041/ijsmt.v2i3.011

Abstract—

Land record digitization is a critical component of modern e-governance systems, enabling transparent land administration and efficient property management. Large-scale cadastral digitization initiatives often face data quality challenges such as boundary overlaps, geometric inconsistencies, and ownership mismatches due to legacy data migration, manual surveying errors, and heterogeneous data sources. Traditional rule-based validation techniques are limited in scalability and effectiveness for nationwide land digitization programs. This paper proposes an intelligent machine learning–based framework for automated anomaly detection in digitized land records, with a case study of the land record digitization platform. The proposed framework integrates spatial, topological, and attribute-based features extracted from cadastral data and applies unsupervised machine learning models to identify anomalous records without requiring labelled training data. Experimental evaluation on real-world digitized land records demonstrates that the proposed approach significantly improves anomaly detection accuracy and reduces manual verification effort compared to conventional validation methods. The study highlights the potential of intelligent data-driven techniques to enhance the reliability and scalability of large-scale land record digitization systems.

Keywords— Land Record Digitization, Cadastral Data, Machine Learning, Anomaly Detection, GIS, E-Governance.

I. INTRODUCTION

Accurate land records form the foundation of land administration, property rights management, taxation, and dispute resolution. Many countries have undertaken large-scale land record digitization initiatives to modernize legacy paper-based cadastral systems. In India, the digital land records platform has been introduced to support systematic cadastral mapping and integration of spatial and ownership data.

Despite these efforts, digitized land records often contain inconsistencies introduced during surveying, digitization, and data integration processes. Common issues include overlapping parcel boundaries, gaps between adjacent parcels, irregular parcel shapes, and

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **03** March-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

mismatches between spatial boundaries and ownership attributes. Manual verification of such anomalies is time-consuming, costly, and not feasible at national scale.

Machine learning (ML) offers a promising solution by enabling automated detection of irregular patterns in large and complex datasets. In particular, unsupervised anomaly detection techniques are well suited for cadastral datasets where labeled examples of errors are scarce. This paper investigates the application of machine learning–based anomaly detection techniques to digitized land records as a real-world case study.

The main contributions of this paper are:

- A machine learning–based framework for automated anomaly detection in digitized land records
- A comprehensive feature engineering strategy combining spatial, topological, and ownership attributes
- Comparative evaluation of unsupervised ML models against traditional rule-based validation
- Practical insights from applying the framework to real cadastral data from the digitize platform.

II. LITERATURE REVIEW

Cadastral data quality and land record digitization have been extensively studied in the fields of GIS and land administration. Traditional validation approaches rely on topology rules, manual inspection, and field verification. While effective for small datasets, these methods struggle with scalability and adaptability.

Recent research has explored machine learning techniques for spatial data quality assessment and anomaly detection. Unsupervised methods such as clustering, isolation-based models, and density-based algorithms have been applied to geospatial datasets including road networks and urban maps. However, limited work has focused specifically on large-scale cadastral land records.

This paper addresses this gap by presenting a machine learning–based anomaly detection framework tailored for digitized land records and validated using operational data from the digitized platform.

III. OVERVIEW

IV. METHODOLOGY

System Architecture

The proposed framework follows a modular, layered architecture designed for integration with existing land digitization systems.

Figure 1. System Architecture of the Proposed Framework

Platform

(Digitized Land Records)

|

Data Ingestion Layer

|

Pre-processing & Validation

|

Feature Engineering

|

ML-Based Anomaly Detection

(Isolation Forest, OC-SVM, DBSCAN)

|

Visualization & Decision Support

Figure 1: illustrates the layered architecture of the proposed intelligent framework for automated anomaly detection in digitized land records.

Data Pre-processing

Pre-processing ensures data consistency and includes:

- Removal of duplicate records
- Correction of invalid geometries
- Coordinate reference system normalization
- Handling missing or inconsistent attribute values

Basic rule-based checks are applied to filter obvious errors before ML processing.

Feature Engineering

Three categories of features are extracted:

Spatial Features: parcel area, perimeter, compactness, and area deviation
Topological Features: overlaps, gaps, and adjacency consistency
Attribute-Based Features: ownership frequency, parcel–ownership consistency, and missing values

All features are normalized to ensure uniform scaling.

Anomaly Detection Models

Unsupervised machine learning models are used due to the lack of labelled data:

- **Isolation Forest:** isolates anomalies using random partitioning
- **One-Class SVM:** learns a boundary around normal data
- **DBSCAN:** identifies low-density points as anomalies

Each model outputs anomaly scores used to flag suspicious records.

Experimental Setup

The framework was implemented using Python with standard GIS and machine learning libraries. Spatial operations were performed using GIS tools, and ML models were implemented using established ML frameworks.

Evaluation Methodology

Detected anomalies were validated through manual inspection and comparison with traditional rule-based validation. Performance was measured using precision, recall, and F1-score.

Table 1. Performance Comparison of Anomaly Detection Methods

Method	Precision	Recall	F1-Score
Rule-Based Validation	0.68	0.62	0.65
DBSCAN	0.74	0.70	0.72
One-Class SVM	0.78	0.73	0.75
Isolation Forest	**0.83**	**0.79**	**0.81**

Isolation Forest achieved the highest overall performance, effectively detecting complex spatial and ownership anomalies missed by rule-based methods.

V. DISCUSSION

The experimental results demonstrate that machine learning–based approaches significantly improve anomaly detection accuracy in digitized land records. The framework reduces manual verification workload by prioritizing high-risk records. While effective, challenges remain in model interpretability and threshold selection, which can be addressed through domain expert feedback.

VI. CONCLUSION AND FUTURE WORK

This paper presented an intelligent machine learning–based framework for automated anomaly detection in digitized land records, validated using the digitized platform as a case study. The proposed approach enhances data quality assurance by detecting spatial and ownership inconsistencies at scale. Future work will focus on hybrid rule-based and ML approaches, deep learning models for spatial pattern recognition, and real-time validation for large-scale land governance systems.

REFERENCES

1.Lemmen, A., et al. "**Land Use Policy**",2012.

2.Van Oosterom, P., et al. "**Computers, Environment and Urban Systems**",2017.

3.Breunig, M., et al. "**ISPRS International Journal of Geo-Information**",2020.

4. Liu, F. T., et al. "In *Proceedings of the IEEE International Conference on Data Mining (ICDM)",* 2008.

5. Schölkopf, B., et al. " **Neural Computation**",2001.

6.Ester, M., et al. "In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)",* 1996.

7. Zhang, J., & Goodchild, M."CRC Press", 2014.

8. Rajabifard, A., et al. " **Land Use Policy"**, 2020.

9. Hastie, T., et al. " Springer",2017.

10.Han, J., et al. "Elsevier", 2012.

11. Goodchild, M. " Wiley",2018.

12. Duckham, S., et al. " In *Proceedings of the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS)",*2008.