



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.935

Volume 15, Issue 10, October 2026

AI-Powered Digital Land Record Management and Land Value Prediction System

**Prof. C. W. Rawarkar, Divya Sandip Gadhawe, Devyani Laxman Raut, Abhishek Rajesh Raut,
Prajwal Sanjay Charbhe**

Guide, Department of Electronic and telecommunications, Prof. Ram Meghe Institute of Technology and Research,
Badnera, Amravati, Maharashtra, India

Student, Department of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research,
Badnera, Amravati, Maharashtra, India

ABSTRACT: Land records, ownership transfers and property valuation are usually managed separately, even after the digitalization drive under DILRMP. This paper proposes an AI-powered platform that combines digital land records, mutation history, GIS-based visualization and QR-code document verification with machine-learning land-value prediction. Land value is estimated from location, area, land type, road connectivity, nearby facilities, transaction history and government ready reckoner rates. Linear Regression, Decision Tree, Random Forest and XGBoost are compared using MAE, RMSE and R^2 . The platform supports transparent administration and informed valuation decisions.

KEYWORDS: Digital Land Records, Land Value Prediction, Machine Learning, GIS, XGBoost.

I. INTRODUCTION

Land is a major economic and social resource, and accurate ownership information is central to property transactions, infrastructure construction, taxation, agriculture planning and public administration. In India, land administration is handled by State and Union Territory governments, which leads to differences in procedures, record formats, terminology and levels of digitization. The Government of India therefore introduced the Digital India Land Records Modernization Programme (DILRMP) to modernize land-record management, improve transparency, reduce land disputes and fraudulent transactions, and widen access to land information [7].

Despite progress in digitizing Records of Rights, cadastral maps and registrations, digitized records alone do not remove every land-administration problem. Information stays scattered across administrative processes, and citizens may need separate procedures to verify records, apply for mutations, retrieve documents and estimate property value. Interoperability, up-to-date data, spatial integration and convenient access remain key requirements [7].

Land-value estimation is a second challenge. Market value depends on interacting factors such as location, size, land-use class, road connectivity, proximity to cities and highways, public infrastructure, past transaction prices and government valuation rates. Manual approaches need significant effort and may give inconsistent results. Structured property and transaction data make this problem suitable for machine learning [1].

Machine learning has already been applied to property-price prediction. Park and Bae compared several approaches on housing data from Fairfax County, Virginia, and showed that models can capture the relationship between property features and price [1]. However, housing prediction and land valuation are not identical, because land valuation involves additional geographical, administrative and land-use characteristics. This paper addresses that gap by proposing one platform that joins digital land-record management with machine-learning-based land-value prediction. The paper is organized as follows. Section II surveys related work. Section III describes the methodology. Section IV compares the proposed system with previous research and presents the evaluation framework. Section V concludes the paper.

II. LITERATURE SURVEY

Research on digital land administration and intelligent property valuation draws on machine learning, GIS, data management and secure digital governance. Existing studies fall into two groups: machine-learning methods for property-price prediction, and technologies that improve the reliability and security of ownership records.

Park and Bae [1] applied machine-learning algorithms to the Fairfax County housing dataset and showed that algorithm comparison matters; their work covered housing only, not agricultural or general land parcels, land records or land administration. Dipanshu et al. [2] studied regression techniques for house-price prediction using structural and topographical features, found that feature selection affects performance, and reported strong results from the Decision Tree Regressor, but did not address land records, mutation history, GIS visualization or document validation. Adetunji et al. [3] used Random Forest for house-price prediction and found that the ensemble approach handles complex relationships well, yet their study ignored land-administration data and GIS-assisted analysis. Chen and Guestrin [8] introduced XGBoost, a scalable tree-boosting system well suited to structured tabular data.

Shuaib et al. [4] proposed a blockchain-based framework for land registries that targets modification, duplication, reliability and trust, but it does not estimate market value or use machine learning. Thakur et al. [5] analyzed blockchain for land records and land titling in India, noting benefits for transparency and reliability alongside legal, organizational and technological barriers. Krishnapriya and Sarath [6] proposed a blockchain model that secures land registration and reduces tampering risk, without land-value estimation, GIS visualization or a valuation service for citizens.

Property-price prediction and digital land registration have each received substantial attention, but they are generally treated separately. No surveyed work combines digital land-record management, mutation history, GIS visualization, document verification and machine-learning land-value prediction on one platform [1], [3]–[6]. This is the research gap addressed here.

III. METHODOLOGY

The proposed system is a single platform that manages digital land records and predicts land value with machine learning. The work follows the nine-stage pipeline in Fig. 1, from data collection to output generation.

System Flow Diagram

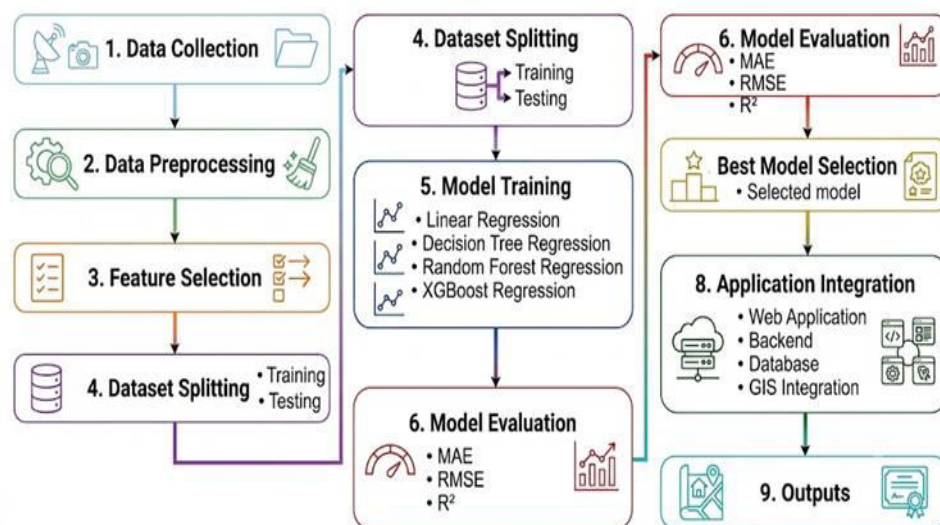


Fig.1. System flow diagram of the proposed platform

A. Land-record management modules

Ownership details, survey number, area, land type, village, taluka, registration information and mutation history are stored in one centralized database. A GIS module, built on Google Maps or OpenStreetMap, shows the parcel and its surrounding infrastructure so that geographical factors affecting value can be examined. A QR-code module lets users verify land documents and open the related digital records.

B. Data collection and preprocessing

Land and transaction data are gathered from web sources by a scraper and extracted from JSON, XML, XLS and PDF files. The data are then cleaned and processed, stored in the database and passed to the prediction model, as shown in Fig. 2 [7].



Fig.2. Data collection and processing pipeline (web scraping to ML model)

C. Feature selection

The valuation model uses location, land area, land type, road connectivity, distance from highways and urban areas, proximity to important infrastructure, past land prices and government-defined ready reckoner rates. Using several factors gives a wider basis for valuation than any single factor.

D. Model training and evaluation

The dataset is split into training and testing sets. Linear Regression, Decision Tree Regression, Random Forest Regression and XGBoost Regression are trained on historical land prices [8]. Each model is tested with Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and the coefficient of determination (R^2), and the best-performing model is selected for deployment. With y_i the actual value, $\hat{y}_i$ the predicted value, $\bar{y}$ the mean actual value and n the number of test samples:

$$MAE = (1/n) \sum |y_i - \hat{y}_i| \quad (1)$$

$$RMSE = \sqrt{(1/n) \sum (y_i - \hat{y}_i)^2} \quad (2)$$

$$R^2 = 1 - \sum (y_i - \hat{y}_i)^2 / \sum (y_i - \bar{y})^2 \quad (3)$$

E. Technology stack

The backend uses Java Spring Boot and REST APIs. The frontend uses React with Redux Toolkit for state management. MySQL and/or PostgreSQL store land records, transaction records, user details, geographic attributes, document references and prediction results. The machine-learning models are built in Python and connected to the web application through APIs or services.

IV. RESULTS & DISCUSSION

The proposed platform covers more of the land-administration and valuation workflow than any single surveyed study. Table I compares it with previous research on input data, technique, output and limitation.

TABLE I. COMPARISON OF PREVIOUS RESEARCH WITH THE PROPOSED SYSTEM

Ref.	Study	Input data	Algorithm technology /	Output	Limitation addressed
[1]	Park & Bae (2015)	Fairfax County housing attributes	Machine-learning regression	Predicted housing prices	Housing only; no land records or GIS
[2]	Dipanshu et al. (2023)	Location, area, property attributes	Regression / ML	Predicted house prices	No mutation history, GIS or document validation
[3]	Adetunji et al. (2022)	Property features	Random Forest	Predicted house prices	No land-administration data or GIS
[4]	Shuaib et al. (2020)	Land tenure records, transactions	Blockchain, distributed ledger	Secure land registry	No value prediction or ML
[5]	Thakur et al. (2020)	Ownership, titles, registration records	Blockchain / DLT	Tamper-proof land records	No intelligent valuation
[6]	Krishnapriya & Sarath (2020)	Registration and ownership data	Blockchain	Secure registration records	No ML valuation or GIS
[7]	DILRMP	Land records, maps, registration data	Digital records, GIS, databases	Computerized land records	Valuation handled separately
[8]	Chen & Guestrin (2016)	Structured tabular data	XGBoost	Accurate, efficient prediction	General method, not applied to land
Prop .	This paper	Location, area, land type, access, price history, ready reckoner rates, nearby facilities	Random Forest, XGBoost, web stack, GIS	Predicted land value and map-based record view	Integrates all of the above

Housing-price studies [1]–[3] show that machine learning captures property features but ignore administrative land data. Blockchain studies [4]–[6] secure ownership records but offer no valuation. DILRMP [7] digitizes records but leaves valuation to separate processes. The proposed system joins these strands and uses XGBoost [8] because it handles structured tabular data efficiently.

Linear Regression, Decision Tree, Random Forest and XGBoost will be compared on the same test set using MAE, RMSE and R^2 . Lower MAE and RMSE and a higher R^2 indicate a better model. Table II is the template for the measured values and is left blank until experiments are run.

TABLE.II. MODEL PERFORMANCE ON THE TEST SET (TO BE COMPLETED)

Model	MAE	RMSE	R^2
Linear Regression			
Decision Tree Regression			
Random Forest Regression			
XGBoost Regression			

The system supports valuation decisions but does not replace official records, legal procedures or authorized valuation. Its accuracy depends on data quality, completeness, security and geographical coverage.

V. CONCLUSION

The proposed system brings digital land administration and machine-learning-based land-value estimation onto one platform. It organizes land records, ownership details, mutation history, documents and geographic information, and gives users a decision-support tool for assessing land value. Comparing Linear Regression, Decision Tree, Random Forest and XGBoost through MAE, RMSE and R^2 gives a systematic way to choose the model that suits the available dataset. Centralized record management improves access, GIS adds geographical context, and QR codes support document verification.

The system is meant to assist administration and valuation decisions, not to replace official government records, legal procedures, ownership verification, mutation approval or authorized valuation. Its effectiveness depends on data quality, completeness, security and geographical coverage, so privacy protection, access control, data integrity, model transparency, bias monitoring and administrative accountability are part of the design [7]. Future work is to train the models on real land-transaction data and report the measured results.

REFERENCES

- [1] B. Park and J. K. Bae, "Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data," *Expert Systems with Applications*, vol. 42, no. 6, pp. 2928–2934, 2015, doi: 10.1016/j.eswa.2014.11.040.
- [2] Dipanshu, Tarun, Harsh, and Keshav, "House Price Prediction Using Machine Learning," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 11, no. 7, 2023, doi: 10.22214/ijraset.2023.54739.
- [3] A. B. Adetunji, O. N. Akande, F. A. Ajala, O. Oyewo, Y. F. Akande, and G. Oluwadara, "House Price Prediction using Random Forest Machine Learning Technique," *Procedia Computer Science*, vol. 199, pp. 806–813, 2022, doi: 10.1016/j.procs.2022.01.100.
- [4] M. Shuaib, S. M. Daud, S. Alam, and W. Z. Khan, "Blockchain-based framework for secure and reliable land registry system," *TELKOMNIKA*, vol. 18, no. 5, pp. 2560–2571, 2020, doi: 10.12928/telkomnika.v18i5.15787.
- [5] V. Thakur, M. N. Doja, Y. K. Dwivedi, T. Ahmad, and G. Khadanga, "Land records on Blockchain for implementation of Land Titling in India," *International Journal of Information Management*, vol. 52, p. 101940, 2020, doi: 10.1016/j.ijinfomgt.2019.04.013.
- [6] S. Krishnapriya and G. Sarath, "Securing Land Registration using Blockchain," *Procedia Computer Science*, vol. 171, pp. 1708–1715, 2020, doi: 10.1016/j.procs.2020.04.183.
- [7] Department of Land Resources, Ministry of Rural Development, Government of India, "Digital India Land Records Modernization Programme (DILRMP)." [Online]. Available: <https://dolr.gov.in/en/programmes-schemes/dilrmp-2/>
- [8] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details