

PREDICTIVE ANALYTICS-BASED DECISION SUPPORT SYSTEM FOR LAND INVESTMENT USING AI

M.Arokia Muthu^{1,*}, A.Rohith Kumar², B.Nithin², B.Akshitha², D.Deethvik², G.Sathwik²

¹Assistant Professor, Department of CSE(DS), TKR College of Engineering, Meerpet, Telangana 500097

²B.Tech (Scholar), TKR College of Engineering, Meerpet, Telangana 500097

*Correspondence: arokyamuthu@tkrcet.com

ABSTRACT

Land investment is considered a promising long-term financial strategy, but evaluating the true potential of land plots is often difficult due to limited analytical tools and heavy reliance on brokers. Most real estate platforms mainly focus on built properties and provide only basic information for raw land evaluation. This research proposes a predictive analytics-based decision support system that utilizes artificial intelligence to assist investors in making informed land investment decisions. The system integrates multiple datasets including historical land price trends, geographic location data, transportation accessibility, environmental indicators, and safety statistics. Machine learning algorithms such as Linear Regression, Random Forest, and K-Means Clustering are applied to analyze patterns and forecast future land value appreciation. A composite Investment Score is calculated by combining predicted growth with civic and environmental factors. Users can filter plots based on budget and visualize results through an interactive map interface. The platform also generates analytical reports summarizing investment potential and comparative insights. The modular architecture allows future expansion to multiple cities and integration with additional data sources. This approach enhances transparency and enables data-driven real estate investment decisions

Keywords: Predictive Analytics, Artificial Intelligence, Land Investment Analysis, Machine Learning, Decision Support System, Real Estate Analytics, Investment Scoring, Geographic Data Analysis, Smart Property Evaluation, Data-Driven Real Estate.

1. INTRODUCTION

The real estate industry has experienced significant growth with the introduction of digital platforms that simplify property search and comparison. These platforms allow users to access property listings, compare prices, and explore different locations through online interfaces. However, most existing systems are

primarily designed for built properties such as apartments and houses. As a result, they provide only limited analytical support for evaluating raw land investments and identifying long-term growth potential. Another limitation of current real estate platforms is the absence of predictive analytics capabilities. Most systems display static information such as location, price, and plot size without estimating how land values may change in the future. Without predictive insights, it becomes difficult for investors to compare multiple land plots or identify regions with strong development potential. Therefore, there is a growing need for intelligent systems that can analyze land investment opportunities more effectively.

To address these challenges, this research proposes a predictive analytics-based decision support system that utilizes artificial intelligence and machine learning techniques. The system integrates datasets such as historical price trends, geographic coordinates, metro accessibility, pollution levels, and crime statistics. By analyzing these parameters, the platform predicts future land value appreciation and calculates a comprehensive Investment Score. This approach enables investors to make transparent, data-driven decisions while exploring potential land investments through interactive visualization tools.

2. METHODOLOGY

The proposed system follows a structured methodology to analyze and rank land plots based on their investment potential. The framework integrates multiple stages including data collection, preprocessing, predictive modeling, investment score computation, and visualization. These stages work together to transform raw land data into meaningful insights that help investors identify promising locations for land investment.

2.1 Data Collection

The first stage involves collecting datasets from multiple reliable sources such as government open data portals,

real estate platforms, and publicly available geographic databases. The collected dataset contains attributes including land price history, latitude and longitude coordinates, distance from metro stations, pollution levels, crime statistics, and infrastructure development indicators. These parameters represent important civic and environmental factors that influence long-term land value appreciation. The collected data forms the foundation for the predictive analysis performed by the system.

2.2 Data Processing

After collecting the raw datasets, preprocessing is performed to ensure data quality and consistency. Missing values are handled through appropriate techniques such as imputation or removal of incomplete records. Data normalization is applied to standardize different attributes so that each parameter contributes fairly during model training. Feature engineering techniques are also used to derive additional indicators such as accessibility scores, environmental scores, and civic development indexes that enhance the analytical capability of the predictive models.

2.3 Machine Learning Models

- **Linear Regression** is applied to predict future land prices based on historical trends and civic parameters, providing interpretable growth forecasts.
- **Random Forest** is employed to capture non-linear relationships among features, improving prediction accuracy and reducing overfitting.
- **K-Means Clustering** groups similar land plots based on multiple attributes such as price, civic score, and predicted growth, highlighting high-potential investment zones.

2.4 METHODOLOGY DIGRAM



Fig.1 Methodology Diagram

2.4 Investment Score Calculation

To provide a comprehensive evaluation of each land plot, the system calculates a composite Investment Score. This score is generated using a weighted scoring mechanism that integrates multiple parameters such as predicted price growth, metro accessibility, crime rate, pollution level, and infrastructure quality. Each factor is assigned a weight based on its relative importance in influencing land value. The resulting Investment Score enables objective comparison and ranking of land plots for investment purposes.

2.5 Visualization and User Interaction

The final stage of the methodology focuses on presenting the analytical results to users in an intuitive and interactive manner. The shortlisted land plots are displayed on an interactive map interface using geospatial visualization tools. Users can input their budget constraints to filter suitable plots and compare their investment scores. Additionally, the system generates downloadable reports summarizing key

metrics, predicted growth trends, and ranking details to support informed decision-making.

3. PROPOSED SYSTEM

The proposed system, LandScope, is an AI-based decision support platform designed to assist investors in evaluating land plots based on their investment potential. Unlike traditional real estate platforms that primarily focus on built properties such as apartments and houses, Land Scope specifically targets raw land analysis by integrating predictive analytics and civic intelligence. The system aims to eliminate dependency on brokers and enable investors to make transparent, data-driven decisions using objective evaluation criteria. The platform analyzes multiple parameters that influence land value, including historical land price trends, proximity to metro stations, crime rates, pollution levels, and infrastructure development indicators. A weighted scoring mechanism combines these parameters to generate a comprehensive Investment Score, which allows users to compare and rank plots based on their overall investment suitability. Users interact with the system through a web-based interface where they can input their investment budget and preferences. Based on the provided budget, the system filters the available land plots and displays only those that meet the affordability criteria. The selected plots are then ranked according to their Investment Scores and presented on an interactive map interface. This visualization enables users to easily identify high-potential areas and analyze location-specific factors. To support deeper analysis, the system also generates downloadable reports summarizing key metrics such as predicted growth, civic indicators, and investment rankings. The modular design of LandScope allows the integration of additional datasets, real-time property listings, and advanced predictive models in future updates. By combining machine learning, geospatial visualization, and data-driven scoring, the proposed system provides a comprehensive framework for intelligent land investment analysis.

4. LITERATURE SURVEY

Recent research has demonstrated the potential of machine learning and deep learning techniques for real estate price prediction and investment analysis. Graph-based models such as Graph Convolutional Networks

(GCNs) have been applied to capture spatial dependencies between properties and surrounding infrastructure.[1] Approaches like the A-SRGCNN model incorporate attention mechanisms to improve valuation accuracy by analyzing spatial and temporal relationships in large metropolitan datasets [2-3]. Similarly, the LUCE framework utilizes lifelong learning with heterogeneous property graphs to continuously update property valuations. These models highlight the importance of integrating spatial and temporal information in real estate analytics [4]. Deep learning methods have also been explored to analyze real estate transaction data and identify patterns influencing property prices [5]. Techniques combining autoencoders, clustering algorithms, and optimization strategies have been used to extract key factors such as transaction volume, unit price, and economic indicators [6-8]. Multi-modal stacking-based ensemble models further enhance prediction performance by integrating multiple data sources, including structural attributes, geographic features, and visual property information [9]. These advanced approaches demonstrate improved predictive accuracy but often require extensive multi-source datasets and complex computational resources [10-11].

In addition to predictive modeling, clustering and geospatial analysis techniques have been used to identify high-value regions in urban environments [12]. Methods such as K-Means clustering and Multi-Reference Clustering (MRC) group geographic data points based on spatial similarities and investment indicators [13-15]. Visualization-driven systems like Waldo enable interactive exploration of large geospatial datasets, helping analysts detect hidden patterns and regional trends [16]. These techniques are particularly useful for identifying potential investment zones but may require careful parameter tuning and large-scale spatial data for optimal performance [17].

Despite these advancements, most research models focus primarily on residential or commercial property valuation rather than raw land investment analysis [18]. Additionally, existing commercial real estate platforms function mainly as listing services and lack predictive analytics, multi-factor scoring, and civic intelligence integration [19]. These limitations highlight the need for an intelligent system that combines machine learning, environmental data, and infrastructure indicators to

evaluate land investment opportunities [20]. The proposed Land Scope platform addresses this gap by integrating predictive analytics, investment scoring, and interactive visualization into a unified decision-support system.

5. RESULTS

The proposed Land Scope system was evaluated using multiple machine learning techniques, including Linear Regression, Random Forest, and K-Means clustering, to assess its effectiveness in land investment analysis. The models were trained on datasets containing historical land prices along with civic and environmental parameters. Among the evaluated methods, Random Forest demonstrated higher predictive performance due to its ability to model complex relationships between multiple features. Linear Regression provided interpretable results but showed comparatively lower accuracy, while K-Means clustering effectively grouped land plots into categories based on investment potential.

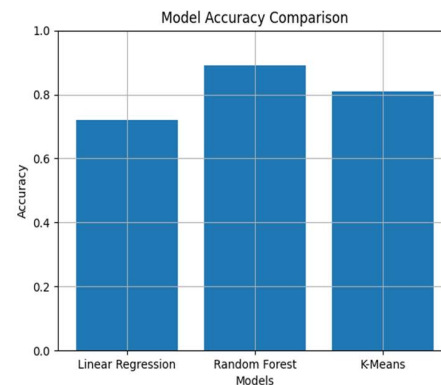
The system successfully predicted future land value trends by analyzing historical data and infrastructure-related factors. The predicted results indicate that areas with better connectivity, lower crime rates, and improved infrastructure tend to exhibit higher growth potential. The Investment Score mechanism enabled ranking of land plots based on multiple criteria, allowing users to identify suitable investment options within their budget. This multi-factor evaluation approach ensures that decisions are not based solely on price but also consider long-term growth and sustainability factors.

Further analysis of model behavior highlights the significance of different features in influencing land value prediction. Historical price trends and infrastructure development were observed to have a strong impact on the predicted outcomes, followed by accessibility and safety indicators. Environmental parameters such as pollution also contributed to the evaluation, although to a lesser extent. These findings validate the effectiveness of incorporating diverse datasets in predictive analytics and demonstrate that the proposed system provides reliable and practical decision support for land investors.

6. IMPLEMENTATION

The Land Scope platform is implemented as a web-based analytical system that integrates machine learning with geospatial visualization to support intelligent land investment decisions. The backend infrastructure is developed using Python along with the Django web framework, which manages data processing, server-side logic, and interaction with predictive models. Various datasets related to land plots in Hyderabad are collected

from public



sources and real estate repositories. These datasets include attributes such as historical land prices, geographic coordinates, distance from metro stations, environmental indicators, crime statistics, and infrastructure development metrics.

7. IMPLEMENTATION DIAGRAM

Before applying machine learning techniques, the collected data is prepared through preprocessing operations to ensure reliability and consistency. Data cleaning procedures remove incomplete or inconsistent entries, while normalization ensures that different attributes contribute proportionally during model training. Predictive analytics is performed using machine learning algorithms such as Linear Regression for estimating future price growth and Random Forest for capturing complex interactions between multiple factors. K-Means clustering is additionally applied to categorize land plots into groups with similar characteristics, helping identify regions with strong investment potential.

The analytical results are presented to users through an interactive web interface designed using

HTML, CSS, and JavaScript technologies. Geospatial visualization tools such as Leaflet or Folium display the filtered land plots on a map with markers representing their investment

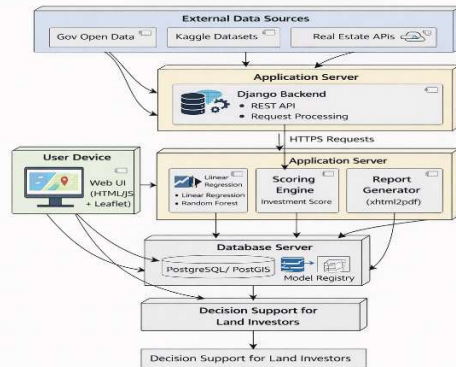


Fig.2 Implementation Diagram

rankings and predicted growth indicators. Users can enter their investment budget to obtain a list of suitable plots along with their computed Investment Scores. The system also generates downloadable PDF reports containing summarized insights, rankings, and location details, enabling users to review the analysis offline. This implementation ensures a scalable and modular architecture that can be extended with additional datasets, real-time property information, and advanced predictive models in future updates.

8. DISCUSSION

The proposed LandScope platform demonstrates the potential of integrating predictive analytics with civic and environmental data to improve land investment decision-making. By utilizing machine learning techniques such as Linear Regression, Random Forest, and clustering algorithms, the system is able to analyze historical trends and multiple location-based parameters to estimate future land value appreciation. This multi-factor evaluation provides investors with deeper insights compared to traditional real estate platforms that only present static information such as price and location. Another important aspect of the system is the integration of interactive visualization tools that simplify the interpretation of complex analytical results. The map-based interface enables users to explore land plots spatially while viewing their investment rankings and predicted growth indicators. Features such as budget-

based filtering and score-based ranking help users quickly identify plots that align with their financial and investment preferences. This improves usability and makes advanced data analysis accessible even for users without technical knowledge of machine learning.

Despite its advantages, the system also faces certain limitations. The accuracy of predictive models depends heavily on the availability and quality of datasets related to land prices, civic infrastructure, crime statistics, and environmental factors. In regions where reliable data sources are limited, prediction performance may be affected. Additionally, the current implementation focuses on a specific geographic region, which may require dataset adjustments when expanding to other cities. Future improvements may include integration with real-time property listings, satellite imagery analysis, and legal verification systems to further enhance the reliability and practical usability of the platform.

9. CONCLUSION

This research presented Land Scope, an AI-driven decision-support platform designed to assist investors in evaluating land investment opportunities using predictive analytics and civic intelligence. The system integrates multiple datasets including historical land price trends, infrastructure accessibility, environmental indicators, and safety metrics to provide a comprehensive evaluation of land plots. By applying machine learning models such as Linear Regression, Random Forest, and K-Means clustering, the platform predicts potential land value growth and computes an investment Score that enables objective comparison between different plots.

The platform enhances transparency and supports data-driven decision-making in the real estate sector. Through features such as budget-based filtering, score-based ranking, and interactive map visualization, users can easily explore and analyze potential investment locations. Additionally, the ability to generate downloadable analytical reports provides users with detailed insights that can be used for further evaluation and planning.

In summary, Land Scope successfully addresses the limitations of existing real estate platforms by combining AI, data science, and civic intelligence into a

unified decision- support framework. It not only provides investors with objective evaluations but also establishes a foundation for future advancements in real estate analytics, contributing to smarter, safer, and more transparent land investment practices.

ACKNOWLEDGEMENTS

We sincerely thank the **Management of TKR College of Engineering and Technology (TKRCET)** for granting us permission and providing the necessary resources and inspiration to carry out this project. Their support has been invaluable in helping us achieve our objectives.

We extend our deepest appreciation to our **Principal, Dr. D. V. Ravi Shankar, M.Tech., Ph.D.**, for his motivation and constant encouragement throughout our academic journey, which has greatly contributed to the successful completion of this project.

We are also thankful to our **Head of the Department, Dr. V. Krishna, M.Tech., Ph.D., Professor, Department of CSE (Data Science), TKRCET**, for his invaluable insights and constructive suggestions, which have helped shape this project.

Finally, we express our heartfelt gratitude to our **Project Coordinator and Internal Guide, Mr. A Muthu, M.E., (Ph.D.), Assistant Professor, Department of CSE (Data Science), TKRCET**, for his continuous guidance, encouragement, and technical expertise, which have played a crucial role in the successful completion of this work.

REFERENCES

1. Z. Hong, R. Zhou, and H. Ai, "A-SRGCNN:A graph convolutional network-based model for megacity real estate valuation," IEEE Access, vol. 10, pp. 104811–104828, 2022.
2. Muthu, M. A., & Prakash, B. (2025). Efficient privacy-preserving mHealth framework using crisscross AES and FCFS-NDPPP in hybrid cloud. *Ingénierie des Systèmes information (ISI)*.
3. Muthu, M. A. (n.d.). Implementation of multi cloud with big data for secured multi purpose smart card authorization using RFID. *International Journal*.
4. Krishna, V., Rajyalakshmi, P., Naresh, P., & Ramesh, V. (2019). A novel IoT-based

authorized accessible and multi-level privacy model for m-healthcare system. *Journal of Xi'an University of Architecture & Technology*, 11(11).

5. Krishna, V., Sumalatha, C., Raju, Y. D. S., & Mohan, K. V. M. (2022). Analysis of heart disease prediction using machine learning classification algorithms. *Journal of Optoelectronics Laser*.
6. Parthiban, L., Latchoumi, T. P., Balamurugan, K., Raja, K., & Parthiban, R. (2023). Cognitive computing for the internet of medical things. In *Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations* (pp. 85-100). Cham: Springer International Publishing.
7. Sneha, N., & Balamurugan, K. (2022, October). Micro-drilling optimization study using RSM on PLA-bronze composite filament printed using FDM. In *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)* (pp. 1-5). IEEE.
8. Pruthviraju, G., Dambhare, S. G., Pathri, B. P., Ramakrishna, M., Gokulanathan, L., Balamurugan, K., & Shumet, W. (2022). Mechanical Test on Aluminum Alloy with Maximal Soluble SiC Reinforcement. *Advances in Materials Science and Engineering*, 2022(1), 9848928.
9. Latchoumi, T. P., Parthiban Balamurugan, K., Raja, K., Vijayaraj, J., & Parthiban, R. (2023). A framework for low energy application devices using blockchain-enabled IoT in WSNs. In *Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations* (pp. 121-132). Cham: Springer International Publishing.
10. Arun Karthikeyan, K., & Balamurugan, K. (2020). Studies on the effects of deep cryogenic treated WC-Co insert on turning of Al6063 using multi-objective optimization. *SN applied Sciences*, 2(12), 2103.
11. Balamurugan, K., Sudhakar, G., Xavier, K. F., Bharathiraja, N., & Kaur, G. (2025). Human-machine interaction in mechanical systems through sensor enabled wearable augmented reality interfaces. *Measurement: Sensors*, 39, 101880 Pavan, M. V., Balamurugan, K., & Balamurugan, P. (2021). Wear experiments on PLA-Cu composite filament printed in different FDM conditions. *Turkish Journal of Computer and Mathematics Education*, 12(9), 2245-2251.

12. Y. Zhong, J. Li, and S. Zhu, "Clustering geospatial data for multiple reference points," *IEEE Access*, vol. 7, pp. 132423–132429, Sept. 2019.
13. D. A. Keim, C. Panse, M. Sips, and S. C. North, "Visual data mining in large geospatial point sets," *IEEE Comput. Graph. Appl.*, vol. 24, no. 5, pp. 36–44, Sept.–Oct. 2004.
14. Krishna, V., Tamrakar, A. K., Banala, R., Saritha, D., Rao, A. L. N., & Buddhi, D. (2022). Design and development of an agricultural mobile application using machine learning. Proceedings of the 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS).
15. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009.
16. Sreenivasa Reddy, K., & Jadhav, P. P. (2023). Integrating AI techniques in 3D virtual worlds through head-mounted display VR systems. *International Journal of Applied Engineering & Technology*, 5(4)
17. Prashanth Kumar, P., & Jadhav, P. P. (2023). Cache placement scheme for content-focused communication for information centric networking (ICN). *European Chemical Bulletin*, 3(1), 3138–3150.
18. Geetha, L. S., El-Ebiary, Y. A. B., Srinivasa Rao, B., Rautrao, R. R., Mastan Rao, T. S., Venkata Naga Ramesh, J., & Al-Omari, O. (2025). Challenges and solutions in agile software development: A managerial perspective on implementation practices. *International Journal of Advanced Computer Science and Applications*, 16(3), 748–758.
19. Venkata Murali Mohan, K., & Krishna, V. (2022, March). Grey hole attack in mobile ad-hoc network mitigation and protection. *IVCMASM 2022 Conference Proceedings*.