Web Based AI Application for Predicting House Construction Cost

Mrs. SHEETAL D N
Pratibha C Y, Ruchin Anandan, Sanjana D, Tanzeela Banu
Information Science and Engineering, Visvesvaraya Technological University.
Sheetal.sowshe@gmail.com
pratibhapreetu28@gmail.com , ruchinanandan11@gmail.com
sanjanafashions2003@gmail.com , tanzeelatanz05@gmail.com

***Abstract**: This project introduces an online AI-powered tool designed to enhance the accuracy and efficiency of estimating residential construction costs. By utilizing machine learning techniques, the system processes historical records, project-specific inputs, and current market data to produce dependable cost estimates. It is capable of managing both structured and unstructured information, capturing the intricate relationships among various cost drivers. The platform is designed to be intuitive, widely accessible via the web, and flexible enough to accommodate diverse construction scenarios. It helps reduce human errors, shortens estimation time, and curtails unexpected budget overruns. It is equipped to handle both structured and unstructured information, allowing it to interpret the intricate factors that influence construction expenses. With an intuitive interface, the tool is accessible from any internet-enabled device and can be adapted to various construction scenarios. Its automation reduces human error, speeds up the estimation process, and helps prevent budget overruns. Ultimately, it aids users in making well-informed financial decisions and streamlining project planning.*

I. INTRODUCTION

Accurately estimating the cost of residential construction is a complex challenge influenced by fluctuating variables such as material availability, labor rates, site location, design intricacies, and market dynamics. Traditional approaches often depend on manual processes, outdated references, or subjective judgment, leading to cost inaccuracies and budget overruns. To overcome these issues, this project introduces a web-based AI platform that employs machine learning algorithms to generate precise and adaptable cost predictions. By analyzing historical data, real-time market inputs, and user-provided project specifications—such as area, number of floors, location, and features like basements or furnishing—the system delivers tailored and data-driven estimates. The online interface is designed for ease of use and broad accessibility, allowing users across devices and backgrounds to efficiently obtain reliable cost insights. The platform continuously updates itself with new data, improving prediction accuracy over time. Serving homeowners, architects, developers, and contractors alike, the tool not only enhances budgeting accuracy but also reduces time, improves transparency, and supports informed decision-making. Ultimately, it brings innovation and intelligence to construction planning, paving the way for a more efficient and technologically advanced approach to cost management.

II. LITERATURE SURVEY

Several studies highlight the transformative role of Artificial Intelligence (AI) and Machine Learning

(ML) in improving construction cost estimation. Saka and Madanayake examined how AI technologies, particularly deep learning, can enhance accuracy in cost prediction, although their application remains limited due to ongoing reliance on traditional methods and data constraints. Similarly, Kamil and Hasan emphasized the growing significance of ML in construction cost analysis and provided a taxonomy of research trends between 2017 and 2021, noting a surge in scholarly interest. Elmousalami's work focused on parametric modeling and hybrid approaches, combining qualitative and quantitative methods with computational intelligence to create more reliable cost models. Finally, Naimi and Salahaldain introduced a Support Vector Machine (SVM)-based model aimed at improving cost estimation during early project phases, offering a shift from expert-based judgment to data-driven predictions. Collectively, these studies support the integration of AI-driven models to increase the efficiency, reliability, and objectivity of construction cost forecasting.

III. PROBLEM STATEMENT

Predicting house construction costs is a major challenge due to fluctuating material prices, labor costs, and project timelines. Factors such as regional labor market dynamics, varying site conditions, and location-specific expenses make accurate cost estimation difficult. The goal of this project is to develop a web-based AI application that leverages historical data and machine learning algorithms to provide accurate cost predictions. The platform will offer users real-time estimates, customizable cost breakdowns, and detailed reports, helping homeowners, contractors, and developers budget effectively. It will incorporate data on labor rates, material prices, and regional factors, making it scalable and adaptable to current market trends. This project aims to create a web-based AI application that uses historical data and machine learning models to provide precise cost predictions for construction projects. By analyzing factors like project size, location, and material quality, the platform will offer real-time cost estimates, detailed breakdowns, and customizable options, helping users better plan and manage their budgets.

IV. BACKGROUND

OpenAI is a leading organization at the forefront of artificial intelligence (AI), with a broad mission to ensure that AI, especially artificial general intelligence (AGI), benefits all of humanity. Its work spans several domains, including natural language processing, computer vision, robotics, healthcare, and education. OpenAI is known for its pioneering advancements in AI, such as the development of GPT (Generative Pre-trained Transformer) and DALL·E, which have set new standards in understanding and generating human-like text and images. The organization's contributions extend beyond technology development to real-world applications. In healthcare, OpenAI's AI models assist in diagnosing diseases and personalizing treatment plans, while in education, they support personalized learning. OpenAI also drives innovation in industries like entertainment and business, with AI tools enhancing creativity and improving customer service. A strong focus on ethical AI deployment ensures that OpenAI's technologies align with human values, addressing safety concerns and promoting transparency. OpenAI is dedicated to making AI accessible to all by providing open-source tools, educational resources, and APIs, empowering users worldwide to benefit from AI technology.

V. SYSTEM REQUIREMENTS

For optimal performance, the proposed application requires a well-defined set of software and hardware specifications. It should operate on Windows 10/11 or Ubuntu 20.04 and above, using Python 3.x and optionally MATLAB for advanced processing. Development tools such as Visual Studio Code, PyCharm, or Jupyter

Notebook are recommended, along with essential libraries like NumPy, Matplotlib, OpenCV, and PyCrypto for handling image analysis, data processing, and encryption. The system's hardware should include at least an Intel Core i5 processor (Core i7 preferred), 8 GB of RAM (16 GB recommended), and 500 GB storage (ideally a 1 TB SSD). Functionally, devices must perform minimal on-site computations to conserve resources. They should capture satellite images automatically, detect water resources, and process images to extract key data independently. These devices should also interact with a central database to synchronize and update data without duplicating previous computations. Scalability and environmental resilience are essential, ensuring reliability across various conditions.

VI. METHODOLOGY

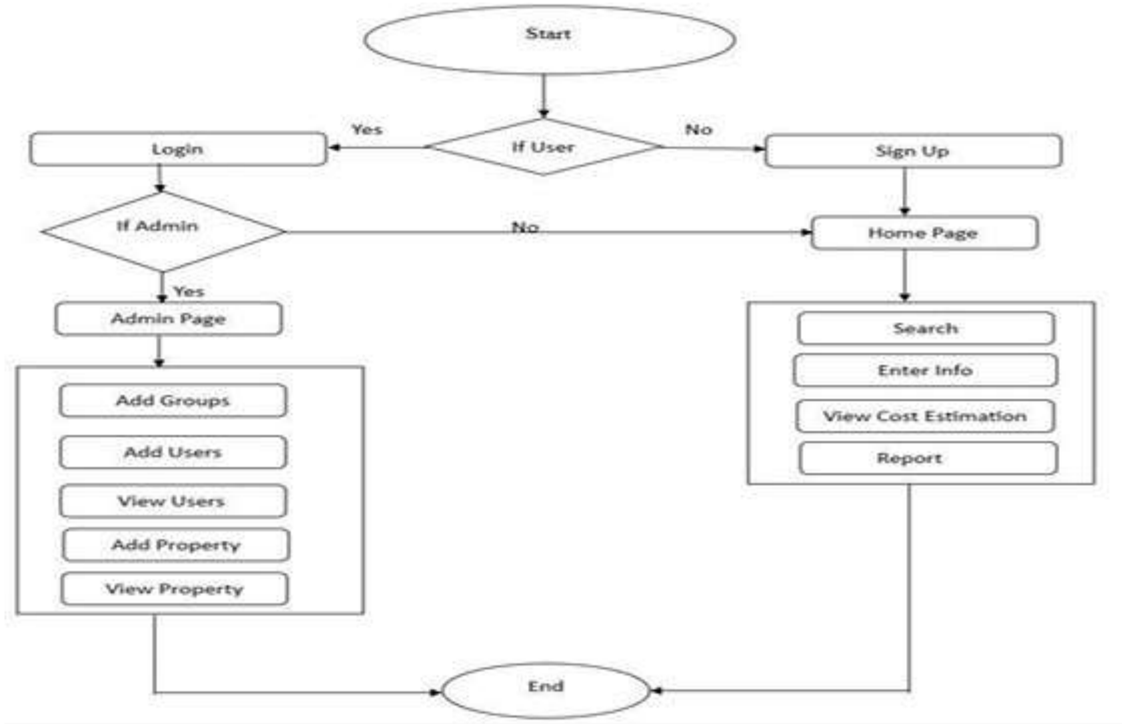


Figure 1: Flow Diagram

The proposed methodology focuses on building a web-based application aimed at predicting house construction costs while offering distinct functionalities for administrators and regular users. The process begins with a user decision point—whether to log in or register. If the user is new, they proceed to the sign-up module. Existing users enter through the login system. After successful login, the system checks the user role. If the user is an administrator, they are redirected to the admin dashboard. Otherwise, they are directed to the user homepage. The admin panel includes features such as adding and viewing property groups, managing users, and inserting or viewing property data, all of which are essential for maintaining and updating the cost estimation system. For standard users, the home page offers access to features such as searching construction templates, entering custom construction parameters, viewing dynamic cost estimations, and generating downloadable reports. The design ensures each user type accesses only the relevant modules to enhance efficiency and usability.The system ensures minimal client-side computation by offloading intensive processing to a centralized server. The flow also incorporates automated image capturing and property analysis, ensuring that the platform is scalable, efficient, and easy to navigate. Once all tasks are completed, the system allows users to safely log out, ending the session.

VII. SYSTEM DESIGN AND DEVELOPMENT

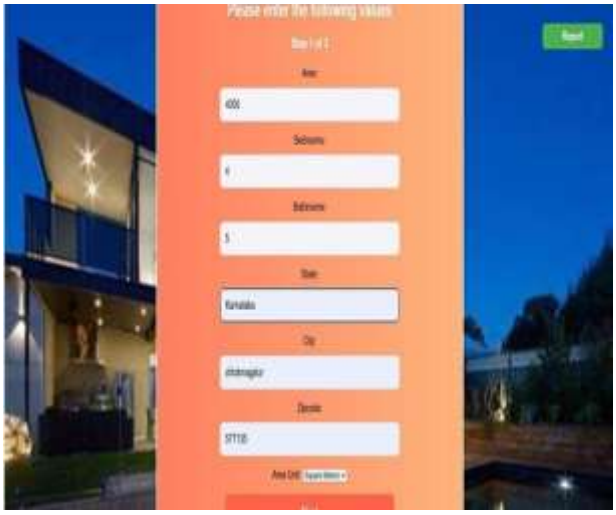
The system architecture for the house construction cost prediction platform is composed of several integrated components to ensure seamless functionality and performance. The frontend, developed using modern frameworks like React, Vue, or Angular, provides features such as user authentication, form-based input for construction parameters, and visual representations of cost predictions through interactive graphs and charts. The backend, powered by technologies like Node.js, Python (Flask/Django), or Java Spring Boot, facilitates data validation, RESTful API communication, and manages interactions with the AI model. The AI/ML module, built using Python libraries such as scikit-learn, TensorFlow, or PyTorch, employs regression techniques (e.g., Random Forest, Linear Regression) and includes preprocessing pipelines, training mechanisms, and deployment capabilities using tools like MLflow. A structured or semi-structured database (PostgreSQL, MySQL, or MongoDB) stores user profiles, historical cost data, and model logs. Deployment and infrastructure leverage Docker, Kubernetes, and cloud services like AWS or GCP, along with CI/CD pipelines (GitHub Actions or Jenkins) to ensure scalability and high availability. The data flow follows a linear path:

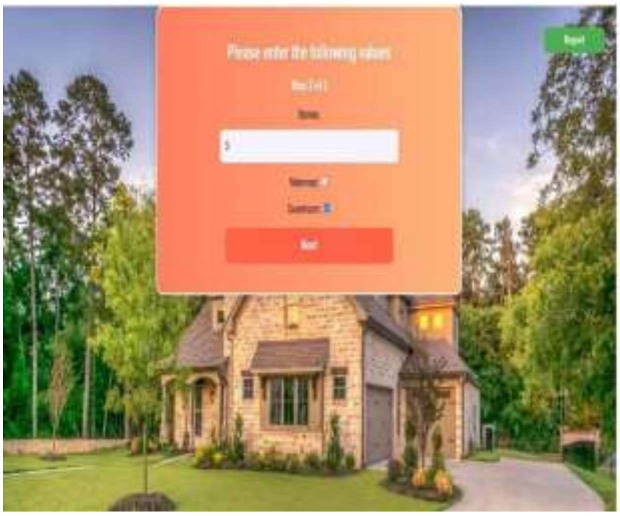
users input data via the frontend, which is processed by the backend and passed to the AI model for inference. Results are returned to the user interface for visualization. The AI model development involves acquiring construction datasets, cleaning and transforming data, engineering features, training models using regression algorithms, and evaluating them with metrics like MAE and RMSE before deployment using APIs such as Flask or FastAPI.

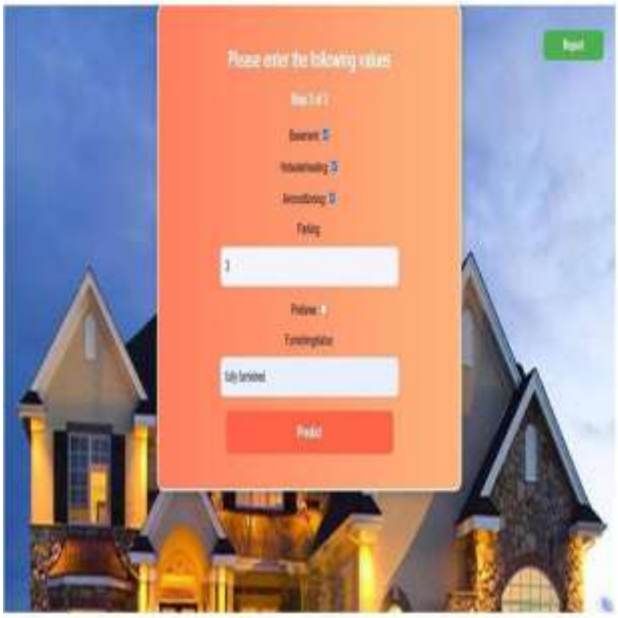
VIII. IMPLEMENTATION

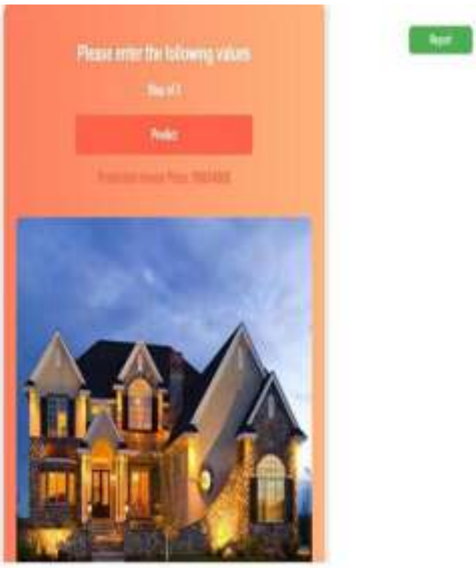
The web-based AI-powered platform for predicting house construction costs is built to offer users an easy-to- navigate system where they can input various property details and receive an estimated cost, excluding material and labor expenses. It emphasizes design-related parameters such as floor area, number of stories, parking space, basement presence, water heating systems, and furnishing levels. The application architecture includes a React- based frontend with Tailwind CSS for responsive design, enabling users to enter data via dropdowns, toggles, and input fields. This data is sent to a FastAPI-powered backend, which validates and preprocesses the inputs before passing them to the AI model. The machine learning model, built using scikit-learn and XGBoost, is trained on historical data, using regression techniques to deliver accurate predictions. It undergoes preprocessing, training, evaluation (using MAE, RMSE, R^2), and is deployed using serialization tools like joblib. The entire system is containerized with Docker for portability and hosted on cloud services such as AWS or GCP to support scalability, load balancing, and seamless deployment.

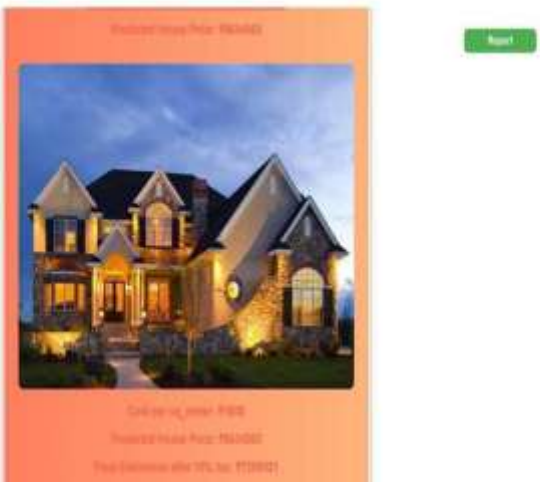
XI. SNAPSHOTS











REPORT

Area	Area unit	Bedrooms	Bathrooms	State	City	Zipcode	Stories	Mainroad	Guestroom	Basement	Hotwaterheating	Air
10000.0	sq_feet	3	2	Karnataka	chikmagalur	577135	2	True	True	True	False	True
12000.0	sq_feet	2	4	Karnataka	chikmagalur	577135	3	True	False	True	False	False
10000.0	sq_meter	2	2	hyderabad	telangana	500213	2	False	True	True	False	True
10000.0	sq_meter	2	2	Karnataka	chikmagalur	577135	2	True	False	True	False	False
4000.0	sq_meter	4	5	Karnataka	chikmagalur	577135	3	False	True	True	True	True
2000.0	sq_meter	2	2	Karnataka	Hassan	573201	2	True	True	True	False	True
3000.0	sq_meter	4	3	Karnataka	Hassan	573201	1	True	False	True	False	True
2000.0	sq_meter	2	2	Karnataka	Hassan	573201	3	False	False	False	False	True
2000.0	sq_meter	3	2	Karnataka	chikmagalur	577135	2	False	True	False	False	False

REPORT

Stories	Mainroad	Guestroom	Basement	Hotwaterheating	Airconditioning	Parking	Prefarea	Furnishingstatus	Predictedprice	Costpersqunit
2	True	True	True	False	True	2	True	semifurnished	9427250	942.72
3	True	False	True	False	False	3	False	fully furnished	7091070	98.49
2	False	True	True	False	True	2	True	fully furnished	6880265	688.03
2	True	False	True	False	False	2	True	fully furnished	6756750	675.68
3	False	True	True	True	True	3	False	fully furnished	6634585	1658.64
2	True	True	True	False	True	2	True	furnished	5975270	2987.64
1	True	False	True	False	True	1	True	unfurnished	5215525	2607.76
3	False	False	False	False	True	2	False	SemiFurnished	4815155	2407.58
2	False	True	False	False	False	2	False	semifurnished	4296526	2148.26

In recent years, the development of web-based AI tools for estimating house construction costs has gained substantial attention. These systems employ advanced machine learning techniques to generate accurate cost forecasts at various stages of a construction project. One notable example is a real estate price prediction tool developed for Bengaluru, India. This application utilized a dataset from Kaggle, where data was thoroughly cleaned and refined to ensure reliability. Important features were extracted to improve model accuracy, and several machine learning models were tested—among which linear regression delivered over 80% accuracy. The tool was deployed using a Flask server, and a user-friendly frontend allowed users to easily interact with the system. In another study focusing on early-stage budgeting, researchers used a natural language processing model known as DistilBERT, which was trained on more than 63,000 building permit records from Victoria, Australia. The model was able to classify residential buildings into different cost categories with 79% accuracy, highlighting the power of AI in handling qualitative data for construction planning. Additionally, hybrid systems combining fuzzy logic, rough set theory, and artificial neural networks have been proposed to enhance prediction reliability. These integrated models help identify critical cost factors and simplify data processing, making them particularly useful during the initial planning phases of construction. Overall, these approaches demonstrate the potential of AI and machine learning in transforming traditional cost estimation into a more precise, data-driven process.

X. CONCLUSION

To sum up, the development of a web-based AI tool for estimating house construction costs represents a major leap forward in making the cost estimation process more efficient and modern. By leveraging advanced machine learning and artificial intelligence, the application provides users—ranging from individual homeowners to contractors and property developers—with precise, real-time estimates based on factors like location, plot dimensions, material choices, and additional building features. This approach replaces outdated manual methods with a quicker, more dependable, and user-friendly alternative. Key functionalities such as detailed cost breakdowns, customizable input options, and downloadable reports help users make informed decisions, enabling improved financial planning and resource management. In essence, the platform connects sophisticated construction analytics with an intuitive interface, offering a powerful tool that minimizes errors, saves time, and enhances the overall experience of predicting construction costs.

XI. REFERENCES

[1] Rad PF (2002) Project estimating and cost management. Management Concepts Incorporated, Virginia

[2] Alsaad HSH (2017) Cost estimation process for construction residential projects by using multifactor linear regression technique. Int J Sci Res 6(6):151–156

[3] Jiang Q (2019) Estimation of construction project building cost by back-propagation neural network. J Eng Des Technol 18(3):601–609

[4] Amin M (2017) Development of cost estimation model for residential building. Int J Civ Struct Eng Res

[5] Sonmez R (2004) Conceptual cost estimation of building projects with regression analysis and neural networks. Can J Civ Eng 31(4):677–683

[6] Byeonghwa Park, Jae Kwon Bae "Using machine learning algorithms for housing priceprediction: The case of Fairfax County, Virginia housing data". 2017

[7] Kuvalekar, Alisha and Manchewar, Shivani and Mahadik, Sidhika and Jawale, Shila, Home cost Forecasting Using Machine Learning (April 8, 2020). Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST)2020.

[8] T. D. Phan, „„Housing price prediction using machine learning algorithms: The case of Melbourne city, Australia,"" in Proc. Int. Conf. Mach. Learn. Data Eng. (iCMLDE), Dec. 2018, pp. 35–42.

[9] L. Yu, C. Jiao, H. Xin, Y. Wang, and K. Wang, „„Prediction on housing price based on deep learning,"" Int. J. Comput. Inf. Eng., vol. 12, no. 2, pp. 90–99, 2018.

[10] Akshay Babu, Dr . Anjana S Chandran, "Literature Review on Real Estate Value Prediction Using Machine Learning", International Journal of Computer Science and Mobile Applications, Vol: 7 Issue: 3, 2019